大 学

数学系

自学丛书

高等代数

·下 册·

GAO DENG DAISHU

大学数学系自学丛书

空间解析几何
高等代数
数学分析
常微分方程
复变函数论
实变函数论
高等几何
计算方法
概率论与数理统计
近世代数
微分几何
电子计算机与算法语言BASIC

统一书号：7090·227
定　　价：2.50 元

大学数学系自学丛书

高 等 代 数

下 册

东北师范大学
贺昌亭 主编

辽 宁 人 民 出 版 社
一九八三年·沈阳

大学数学系自学丛书

高等代数

（下）

贺昌亭　主编

*

辽宁人民出版社出版
（沈阳市南京街6段1里2号）
辽宁省新华书店发行
沈阳新华印刷厂印刷

*

开本：850×1168 1/32　印张：19½
字数：520,000　印数：1-16,200
1983年7月第1版　1983年7月第1次印刷

统一书号：7090·227　定价：2.50元

责任编辑：于　乙
封面设计：安今生

目　　录

练习与习题解答

第七章　集合与映射

通过以上几章的学习，我们可以初步地看到高等代数讨论问题——提出问题、分析问题和解决问题的一些基本特点.

着眼于事物的总体，从事物的总体考虑问题和提出问题.

例如，在初等代数讨论了解二、三元线性方程组的基础上，我们自然地着眼于未知数的个数与方程的个数都不受限制的所有的线性方程组的总体，从线性方程组的总体考虑问题，提出一般线性方程组的求解问题.

对于给定的一个线性方程组来说，我们也是着眼于它的所有的解，从解的总体考虑问题，提出解的结构问题.

再如，在学过一元二次和三次多项式的因式分解问题的基础上，我们自然地着眼于次数不受限制的所有一元多项式的总体，从任意的一元多项式的总体考虑问题，提出一元多项式的整除性问题.

类似地，在学了一元一次和一元二次方程的解法之后，也自然的要考虑一元高次方程的求解问题，提出一元方程的根式解问题.

又如，在第一章关于数的讨论中常常也是这样提出问题的. 由于数数的需要，我们认识了个别的自然数，进而着眼于所有自然数的总体，从自然数的总体考虑问题，提出自然数的结构和运算的封闭性等问题. 对整数、有理数、实数和复数的讨论也是（或者也可以是）这样提出问题的.

从以上说明中可以看出，着眼于事物的总体，从事物的总体考虑问题和提出问题是高等代数的一个重要特点.

着眼于事物的变化和事物间的相互关联，从事物的变化和相互

关联中寻求分析问题的门径，找出解决问题的线索和方法。

例如，线性方程组的求解问题，我们从线性方程组的变形中得到启发，分析同解变换的规律，从这种“形变解不变”的规律中找出解决问题的线索和方法。

对于线性方程组解的结构问题，我们也是从齐次线性方程组与非齐次线性方程组解的变化和相互关联中得到启发，分析解的变化与相互关联的规律，从这种规律中找出解决问题的线索和方法。

再如，把线性方程组抽象为矩阵，对矩阵的行做初等变换来解线性方程组的问题，我们通过矩阵与行列式之间的联系，从这种由矩阵到行列式的变化和相互关联中得到启发，分析矩阵在初等变换下的规律，从“形变秩不变”的规律中找出解决问题的线索和方法。

又如，一元多项式的整除性问题，我们从由多项式到因式、公因式和最大公因式的变化和相互关联中得到启发，分析因式、公因式和最大公因式的规律，从这种规律中找出解决问题的线索和方法。

类似的，对一元高次方程的根式解问题，我们也是从方程的变形和相互关联中找出解决问题的线索和方法。

从以上说明中可以看出，着眼于事物的变化和事物间的相互关联，从事物的变化和相互关联中寻求分析问题的门径，找出解决问题的线索和方法是高等代数的又一重要特点。

运算是讨论问题的基础又是讨论的基本课题，这也是高等代数的一个显著的特点。

例如数的四则运算既是讨论数的基础也是数的讨论的基本课题；而对于多项式，则是在数的四则运算基础上来研究多项式运算的基本性质的；又如 n 维向量空间就是以数的运算为基础来研究数组的一些运算性质的。

综上所述，我们可以初步地看到，着眼于事物的总体，从事物的总体考虑问题和提出问题；着眼于事物的变化和相互关联，从事

物的变化和相互关联中寻求分析问题的门径，找出解决问题的线索和方法；以运算为讨论的基本课题是高等代数的三个重要特点。

一般地说从事物的总体上和事物的变化以及相互关联中掌握事物是近代数学的普遍特点，而以运算作为讨论的基本课题则是代数学区别于其它数学学科的显著特征。

本章在前几章的基础上，初步介绍抽象代数中一些最普遍和最基本的概念和方法，为以后几章的讨论做好必要的准备。同时，对我们系统掌握和深入理解以前几章的内容和方法也是有启发的。

§1 集 合

在第一章我们曾经提到过集合——任意一些事物的总体——的概念。因此，对于集合这一数学用语我们并不陌生，由于从事物的总体上考虑问题和提出问题是高等代数的一个重要特点，于是，我们把事物的总体这样一个概念明确起来，突出出来是非常必要的。这在数学上就是用集合这一术语来表达的。即任何一些事物做为一个总（整）体就叫做一个集合。在数学上把集合当做最基本、最原始的概念，无须另做其它说明。其实，的确也再没有比这更为简单的解释了。

我们已经知道集合、集合的相等、元素、子集等概念及有关的术语和符号。特别地，空集ϕ是有实际意义的，它是客观现象的数学反映。例如，我们用通解这一名称表示一个线性方程组的一切解组成的集合。当所考虑的线性方程组含有n个未知数时，它的通解就是一些n维向量组成的集合。如果该方程组没有解，这时还用通解来表示它的一切解所组成的集合的话，那么这个解集合里就不含有任何n维向量。显然把这样的集合叫做空集是恰当的，也是必要的。并且按子集的定义，我们应当认为空集ϕ是任一集合A的子集，即$\phi\subseteq A$。

本节在上述的基础上，介绍并集、交集、差集、补集、积集、

幂集这样一些基本概念。

设A，B为任二集合。

并集 由属于集合A或属于集合B的一切元素组成的集合叫做A与B的并集，记作

$$A\cup B,\ 即A\cup B=\{x|\ x\in A\ \text{或}\ x\in B\}.$$

交集 由属于集合A同时又属于集合B的一切元素组成的集合叫做A与B的交集，记作

$$A\cap B,\ 即A\cap B=\{x|\ x\in A\ \text{与}\ x\in B\}.$$

如果$A\cap B=\phi$，则称A与B是不相交的。

差集 由属于集A但不属于集B的一切元素组成的集合叫做A与B的差集，记作

$$A-B,\ 即A-B=\{x|\ x\in A但x\overline{\in}B\}.$$

补集 当$B\subseteq A$时，我们把差集$A-B$叫做B在A中的补集，记作B'_A，即$B'_A=\{x|\ x\in A但\ x\overline{\in}B,\ B\subseteq A\}$。

并集、交集可以自然地推广到任意有限个集合的情形。

设$A_1, A_2, \cdots, A_n$是任意n个集合。A_1，A_2，$\cdots$，A_n的并集规定为

$$A_1\cup A_2\cup\cdots\cup A_n=\{x|\ x\in 某一个\ A_i\};$$

$A_1, A_2, \cdots, A_n$的交集规定为

$$A_1\cap A_2\cap\cdots\cap A_n=\{x|\ x\in 每一个\ A_i\}.$$

$A_1, A_2, \cdots, A_n$的并集，交集可以简记为

$$\bigcup_{i=1}^{n} A_i=A_1\cup A_2\cup\cdots\cup A_n;$$

$$\bigcap_{i=1}^{n} A_i=A_1\cap A_2\cap\cdots\cap A_n.$$

下面看几个例子。

例 1 令 $A=\{-1,\ 1\}$，$B=\{0,\ 1\}$，

于是

$$A\cup B=\{-1,\ 0,\ 1\},\quad A\cap B=\{1\},$$

$$A-B=\{-1\},\quad B-A=\{0\}.$$

例 2 考虑 $D^{(2)}=\{(x,\ y)\mid x,\ y\in D\}$的两个子集：

$$X=\{(x,0)\mid x\in D\},\quad Y=\{(0,y)\mid y\in D\},$$

于是

$$X\cup Y=\{(x,y)\mid xy=0\},\quad X\cap Y=\{(0,0)\},$$

$$X-Y=\{(x,0)\mid x\neq 0\},\quad Y-X=\{(0,y)\mid y\neq 0\},$$

$$X'_{D^{(2)}}=\{(x,y)\mid y\neq 0\},\quad Y'_{D^{(2)}}=\{(x,y)\mid x\neq 0\}.$$

例 3 考虑整数集

$$Z=\{\cdots,-3,-2,-1,0,1,2,3,\cdots\}$$

的两个子集：

$$Z_0=\{n\mid 2\backslash n\},$$

$$Z_1=\{n\mid 2\nmid n\}.$$

于是

$$Z_0\cap Z_1=\phi,$$

$$Z_0\cup Z_1=Z.$$

一般地，任意给定一个正整数m，都可以确定整数集 Z 的m 个子集如下：

$$Z_0=\{n\mid n=mq\},$$

$$Z_1=\{n\mid n=mq+1\},$$

$$\vdots$$

$$Z_{m-1}=\{n\mid n=mq+(m-1)\}.$$

此处m个子集的组成情况是清楚的，$n\in Z_r$当且仅当以m除 n 所得余数为r. 因此常常把Z_r叫做以m为模的剩余类，$r=0,1,2,\cdots,m-1$. 容易看出，整数集合Z的这样得到的m个子集具有以下性质：

1) $Z_i\cap Z_j=\phi$，$i\neq j$；

2) $Z=Z_0\cup Z_1\cup\cdots\cup Z_{m-1}$.

例 4 令$M_n(F)=\{(a_{ij})\mid a_{ij}\in F;\ i,j=1,2,\cdots,n\}$. 此处$F$为任意一个数域。于是$M_n(F)$就是$F$上的一切$n$阶方阵组成的集合。考虑$M_n(F)$的如下的一些子集：

$$K_r=\{(a_{ij})\in M_n(F)\mid \operatorname{rank}(a_{ij})=r\};\quad r=0,\ 1,\ 2,\ \cdots,n.$$

这样我们得到$M_n(F)$的$n+1$个子集 K_0，K_1，K_2，…，K_n. 它们的组成情况是清楚的：$(a_{ij})\in K_r$ 当且仅当(a_{ij})的秩等于r. 容易看出：F上n阶方阵的集合$M_n(F)$ 的这样得到的$n+1$个子集有以下性质：

1）$K_i\cap K_j=\phi$，$i\neq j$；

2）$M_n(F)=K_0\cup K_1\cup K_2\cup\cdots\cup K_n$.

例 5 试证：$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$.

证明 按集合相等的定义往证等号两端的集合互为子集即可.

先证 $A\cap(B\cup C)\subseteq(A\cap B)\cup(A\cap C)$. 设$a\in A\cap(B\cup C)$. 由交集的定义，$a\in A$与$a\in B\cup C$. 再由并集定义，从而，$a\in A$，$a\in B$或 $a\in A$，$a\in C$. 即 $a\in A\cap B$ 或 $a\in A\cap C$ 又按并集定义，则得$a\in(A\cap B)\cup(A\cap C)$. 所以 $A\cap(B\cup C)\subseteq(A\cap B)\cup(A\cap C)$.

再证 $(A\cap B)\cup(A\cap C)\subseteq A\cap(B\cup C)$. 设 $a\in(A\cap B)\cup(A\cap C)$. 于是$a\in A\cap B$或$a\in A\cap C$. 亦即$a\in A$，$a\in B$或$a\in A$，$a\in C$. 这等于说：$a\in A$ 同时又有$a\in B$或$a\in C$，即 $a\in A$ 与$a\in B\cup C$. 故有$a\in A\cap(B\cup C)$. 所以又得$(A\cap B)\cup(A\cap C)\subseteq A\cap(B\cup C)$. 总之，$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$.

下面介绍积集、幂集的概念.

设$A_1,A_2,\cdots,A_n$ 为n个非空集合. 其中可以有一些是相同的集合. 任意取定$x_1\in A_1$，$x_2\in A_2$，…，$x_n\in A_n$，把这n个元素组合在一起，记作$(x_1,\ x_2,\ \cdots,\ x_n)$，称其为由 A_1，A_2，…，A_n 所确定的一个n**元组**（也叫 A_1，A_2，…，A_n上的一个n元组）. 两个这样的n元组：

$$\alpha=(x_1,\ x_2,\ \cdots,\ x_n)\text{与 }\beta=(y_1,\ y_2,\cdots,y_n).$$

叫做是相等的当且仅当 $x_1=y_1$，$x_2=y_2$，…，$x_n=y_n$，记作：$\alpha=\beta$.

积集 由 $A_1,A_2,\cdots,A_n$ 所确定的一切 n 元组做成的集合叫做$A_1,A_2,\cdots,A_n$的积集. 记作

$$A_1\times A_2\times\cdots\times A_n.$$

即

$$A_1\times A_2\times\cdots\times A_n=\{(x_1,x_2,\cdots,x_n)\mid x_i\in A_i\}.$$

特别地，当 $A_1=A_2=\cdots=A_n=A$ 时，我们就把这个积集简记为 $A^{(n)}$，其中的元素就叫做A上的n元组。

例 6　令$A=\{0,1\}$，$B=\{-1, 0, 1\}$.

于是

$$A\times B=\{(0,-1),(0,0),(0,1),(1,-1),(1,0),(1,1)\},$$

$$B\times A=\{(-1,0),(-1,1),(0,0),(0,1),(1,0),(1,1)\}.$$

特别的

$$A^{(2)}=\{(0,0),(0,1),(1,0),(1,1)\}.$$

例 7　D为实数域，那么积集

$$D^{(n)}=\{(a_1,a_2,\cdots,a_n)|\ a_i\in D\}$$

就是熟知的实数域D上的n数组集合。当然，这个积集也可以看成是D上的$1\times n$矩阵做成的集合。

最后我们说一下什么叫幂集。设A为任一集合。考虑A的一切子集，于是以A的一切子集为元素又做成一个集合：

$$P(A)=\{X|\ X\subseteq A\},$$

我们把$P(A)$叫做集A的幂集。

例 8　令$A=\{0,1,2\}$，$B=\{0\}$

于是集A共有八个子集如下：

一个元素的子集　$\{0\},\{1\},\{2\}$；

两个元素的子集　$\{0,1\},\{0,2\},\{1,2\}$；

三个元素的子集　$\{0,1,2\}$，即A本身。

不含元素的子集　ϕ，即空集。

这样 A 的幂集 $P(A)$ 就是以上述八个子集为元素做成的集合，即

$$P(A)=\{\{0\},\{1\},\{2\},\{0,1\},\{0,2\},\{1,2\},\{0,1,2\},\phi\}.$$

集 $B=\{0\}$ 只含一个元素，它恰有两个子集，即空集ϕ和它本身。这样B的幂集就是

$$P(B)=\{\phi,\{0\}\}.$$

练 习 一

1. 证明

1) $A\cap B=A$ 当且仅当 $A\subseteq B$;

2) $A\cup B=B$ 当且仅为 $A\subseteq B$,

2. 指出：对集A的任一子集B，存在集C，D使 $B\cup C=A$，$B\cap D=\phi$. 这样的C与D是不是唯一的？如果要求$C=D$时又怎样？

3. 举例说明，有集合A, B使

$A\subset B$，$A=B$，$B\subset A$

三者均不成立.

4. 设 $A=\{a, b, c\}$. 写出 $P(A)$.

§2 映 射

任何事物，不论是个别的或总体的，都不是孤立的、静止的，它的内部和相互之间无不存在一定的联系. 不了解这种联系就不了解事物的规律性. 因此，分析事物的相互联系从中找出规律性的东西，是人们认识事物的必由之路，也是分析问题，解决问题的基本方法. 如前所述，着眼于事物的变化和相互关联，从事物的变化和相互关联中寻求分析问题的门径，找出解决问题的线索和方法也是高等代数的一个重要特点. 这种方法在数学上的体现，就是把所要研究的对象——这里我们把它叫做集合，对它的元素进行比较，从中发现其间的某些联系. 通过这些联系来获得一些对于我们所探讨的问题有益的信息. 这就是本节所要讨论的课题——集合的映射.

设A, B为任二非空集合.

定义1 **对集合A与B给定的一种规则（方法）φ，使得A中每一元素a按着给定的规则φ能在B中确定唯一的一个元素 a'，记作 $a'=\varphi(a)$，这样就说（规则）φ是A到B的一个映射.**

我们用以下符号来表示φ是A到B的一个映射：

$$\varphi: A \rightarrowtail B \text{或} A \overset{\varphi}{\rightarrowtail} B.$$

若 $\varphi: A \rightarrowtail B$，而$a' = \varphi(a)$，我们就说$a'$是$a$（在$\varphi$之下）的象，$a$ 是 a'（在 φ 之下）的一个原象。这时也说φ 把a射成a'，记作：$\varphi: a \longmapsto a'$。

按定义，A中每一元素a都有象 $a' \in B$，而且只有这么一个象；反过来，对B中每一元素x'未必都有原象$x \in A$，如果某一元素x' 有原象的话，也可能不只一个。

设φ_1, φ_2都是A到B的映射。如果对任一元素$a \in A$都有$\varphi_1(a) = \varphi_2(a)$，则称$\varphi_1$与$\varphi_2$是相同的，记作$\varphi_1 = \varphi_2$。

定义 2　设φ是A到B的映射。如果对于B中每一元素（在φ之下）都有原象，则称φ是A到B的满映射。记作

$$\varphi: A \longrightarrow B \text{ 或 } A \overset{\varphi}{\longrightarrow} B.$$

定义 3　设 φ 是A到B的映射。如果对于A中任二不同的元素（在φ之下）的象也不同，则称φ是A到B的单映射，记作

$$\varphi: A \rightarrowtail B \text{ 或 } A \overset{\varphi}{\rightarrowtail} B.$$

如果A到B的映射φ既是单的又是满的，那么就简称φ是A到B的单满映射。记作

$$\varphi: A \rightarrowtail\!\!\!\rightarrow B \text{ 或 } A \overset{\varphi}{\rightarrowtail\!\!\!\rightarrow} B.$$

这时明显地A与B的元素之间一个对一个的对应起来。

A到B能有单满映射φ是一种很特殊而又相当重要的情形。因为我们可以通过A到B的这个单满映射φ唯一地确定一个B到A的映射ψ，并且ψ也是单满的。

事实上　我们利用φ，如下的规定一个关于B到A的规则

$$\psi: B \rightarrowtail A,$$

使得

$$\psi: x' \longmapsto x, \text{ 如果 } \varphi: x \longmapsto x', \ \forall x' \in B.$$

这样，我们首先指出这个规则 ψ 是 B 到 A 的一个映射. 为此只须说明两点：ψ 对 B 的每一元素 x' 都规定了作用，这一点由 φ 是满映射得到保证；ψ 把 B 中每一个元素都射成为 A 中唯一确定的元素， 这是因为 φ 是单映射的缘故.

其次，ψ 也是 B 到 A 的单满映射. 这完全由 φ 是映射所决定.

再次，由 φ 如此确定的单满映射 ψ，明显的只有一个.

由 A 到 B 的单满映射 φ，如上所确定的 B 到 A 的单满映射 ψ 叫做 φ 的**逆映射**. 于是 A 到 B 的任意一个单满映射都有唯一的逆映射 . 这样由 φ 所确定的 B 到 A 的单满映射 ψ 也有自己的逆映射. 显然，ψ 的逆映射恰好是由 A 到 B 的单满映射 φ. 有时也把 A 到 B 的单满映射 φ 叫做 A 与 B 间的一个**一一对应**.

做映射的两个集合 A 与 B 可以是同一个集合，即 $A=B$. 这种特殊情形还是很重要的. 我们把 A 到 A 的一个映射叫做 A 的一个**变换**. 于是也有单变换、满变换、单满变换和逆变换等相应的概念. 因为它们都是非常明白的，所以我们就不一一叙述了. 特别地，A 的单满变换也叫 A 的**一一变换**.

下面通过例子来熟悉上述的一些概念.

例1 令

$$A=\{\cdots,-6,-4,-2,0,2,4,6,\cdots\},$$
$$B=\{\cdots,-3,-2,-1,0,1,2,3,\cdots\}.$$

规则

$$\varphi_1: 2n\longmapsto n,\quad n\text{为任意整数}.$$

这个 φ_1 就是把每一个偶数 $2n$ 射成整数 n. 容易看出，规则 φ_1 是 A 到 B 的一个映射，并且是单满映射，即

$$\varphi_1: A\rightarrowtail\!\!\!\twoheadrightarrow B.$$

因此，在单满映射 φ_1 之下，A 与 B 的元素间一个对一个的对应起来。

规则

$$\varphi_2: 2n\longmapsto 2n,$$

这个φ_2就是把每一个偶数$2n$都射成$2n$本身。容易看出，规则φ_2是A到B的一个映射。并且是单映射，但不是满映射，即

$$\varphi_2: A \rightarrowtail B.$$

规则

$$\varphi_3: 4n \longmapsto n, \quad n\text{为任意整数;}$$
$$2n \longmapsto 2n, \quad n\text{为任意奇数，即 } 2\nmid n.$$

这个φ_3就是把被 4 整除的偶数m——即可以写成$m=4n$，射成用 4 除m的商——即$\frac{m}{4}=n$；而把不被 4 整除的偶数：$2n$，但$2\nmid n$，都射成$2n$本身，容易看出，规则φ_3是A到B的一个映射，并且是满映射，但不是单映射，即

$$\varphi_3: A \twoheadrightarrow B.$$

规则

$$\varphi_4: 2n \longmapsto |2n|.$$

这个φ_4就是把每一个偶数$2n$都射成它的绝对值$|2n|$。容易看出，规则φ_4是A到B的一个映射，它既不是单映射也不是满映射，即

$$\varphi_4: A \rightarrow B.$$

以上所述表明，对给定的两个集合A, B来说，可能建立从A到B的各种可能类型的映射。

φ_3是有启发的，它说明偶数的一部分就能与全体整数一个对一个的对应起来，由此看来，不好说偶数仅仅是整数的一半那么多。

例 2　考虑实数域D上的一切二维向量做成的集合$D^{(2)}=\{(x, y)| x,y\in D\}$。我们对$D^{(2)}$与$D$规定一个规则

$$\varphi_1: (x,y) \longmapsto x.$$

这个φ_1就是把二维向量(x,y)射成它的第一个分量x。容易知道，规则φ_1是$D^{(2)}$到D的一个映射，即

$$\varphi_1: D^{(2)} \rightarrow D.$$

显然，φ_1是满映射，但不是单映射，即

$$\varphi_1: D^{(2)} \twoheadrightarrow D.$$

再给出一个规则

$$\varphi_2:(x,y)\longmapsto\sqrt{x^2+y^2}.$$

这个 φ_2 就是把二维向量 (x,y) 射成它的分量的平方和的算术根 $\sqrt{x^2+y^2}$. 容易知道，规则φ_2是$D^{(2)}$到D的一个映射，即

$$\varphi_2:D^{(2)}\rightarrowtail D.$$

显然，φ_2既不是满映射也不是单映射。

例 3　考虑有理数域Q上m行n列矩阵组成的集合 $M_{m,n}(Q)=\{(a_{ij})|\ a_{ij}\in Q;\ i=1,2,\cdots,m,j=1,2,\cdots,n\}$与有限集合$R=\{0,\ 1,2,\ \cdots,\ m\}$。我们对 $M_{m,n}(Q)$与R给出一个规则

$$\varphi:(a_{ij})\longmapsto \text{rank}(a_{ij}).$$

这个规则φ就是把$m\times n$矩阵(a_{ij})射成它的秩数 $\text{rank}(a_{ij})$. 容易知道，φ是 $M_{m,n}(Q)$到R的一个映射，即

$$\varphi:M_{m,n}(Q)\rightarrowtail R.$$

显然，φ不是单映射。当$m>n$时，φ也不是满映射，当$m\leqslant n$时，φ是满映射，即

$$\varphi:M_{m,n}(Q)\longrightarrow R,\ \text{其中}m\leqslant n.$$

例 4　考虑实数域D上n阶方阵组成的集合$M_n(D)=\{(a_{ij})\ |\ a_{ij}\in D;\ i,\ j=1,2,\cdots,n\}$。我们对 $M_n(D)$ 与实数集D给出一个规则

$$\varphi:(a_{ij})\longmapsto det(a_{ij})=|a_{ij}|.$$

这个规则φ就是把n阶方阵 (a_{ij})射成它的行列式$|a_{ij}|$. 容易知道，φ是$M_n(D)$到D的一个映射，即

$$\varphi:M_n(D)\rightarrowtail D.$$

显然，φ是满映射，但不是单映射（除非$n=1$）。

例 5　令 $A=\{x|\ a\leqslant x\leqslant b;\ \ a,b\in D\}$，即$A$是以$a$，$b$为端点的闭区间上一切实数组成的集合。也可简记作 $A=[a,\ b]$。我们考虑以$[a,\ b]$为定义域的任意一个函数 $y=f(x)$。这样，用f 表示的这个函数规则就是集合 $A=[a,b]$到D的一个映射，即

$$f:[a,b]\rightarrowtail D,$$

f 把闭区间 $[a,b]$ 上每一个实数x射成它的函数值 $y=f(x)$，即

$$f: x \longmapsto y,\ y=f(x).$$

这说明数学分析中讨论的函数都是些特殊的映射，做映射的两个集合都是实数集。

下面再举几个变换的例子。

例 6　考虑 $D^{(2)}=\{(x,y)\mid x,y\in D\}$。令

$\varphi_1:(x,y)\longmapsto(x,0)$；

$\varphi_2:(x,y)\longmapsto\lambda(x,y)$，$\lambda$是一个取定的实数；

$\varphi_3:(x,y)\longmapsto(-x,y)$；

$\varphi_4:(x,y)\longmapsto(x+a,\ y+b)$，$a$ 与b 是取定的两个实数。

容易知道，上述对$D^{(2)}$给出的四种规则都是 $D^{(2)}$ 的变换。其中φ_1既不是单变换也不是满变换；而 φ_3、φ_4都是单满变换，即一一变换。当$\lambda\neq0$时，φ_2也是一一变换。我们已知，任意单满映射都有唯一的逆映射。特别地，这里的φ_3、φ_4、$\varphi_2(\lambda\neq0)$都有逆变换，即

如果用ψ_3表示φ_3的逆变换，那么应有

$$\psi_3:(-x,y)\longmapsto(x,\ y),$$

其实也就是

$$\psi_3:(u,v)\longmapsto(-u,v),$$

即 $\psi_3=\varphi_3$。这表明一一变换φ_3的逆变换就是φ_3自己。

如果用ψ_4表示φ_4的逆变换，那么应有

$$\psi_4:(x+a,\ y+b)\longmapsto(x,y),$$

其实也就是

$$\psi_4:(u,v)\longmapsto(u-a,v-b).$$

如果当$\lambda\neq0$时用ψ_2表示φ_2的逆变换，那么应有

$$\psi_2:(\lambda x,\lambda y)\longmapsto(x,y),$$

其实也就是

$$\psi_2:(u,v)\longmapsto\frac{1}{\lambda}(u,v).$$

上述这四个变换都有一定的几何意义。如果我们把$D^{(2)}$ 看成为建立了笛卡儿直角坐标系的平面上全体点组成的集合，即 (x,y)是在该

坐标系下坐标为 x,y的点。那么φ_1就是往 x 轴上投影的投影变换，也叫射影变换；φ_2就是以λ为系数的相似变换，也叫位似变换；φ_3就是以 y 轴为对称轴的对称变换，也叫反射变换；φ_4就是以 (a,b) 为位移向量的位移变换，也叫平移变换。

例 7　考虑$M_{m,n}(F)=\{(a_{ij})\mid a_{ij}\in F;\ i=1,2,\cdots,m,\ j=1,2,\cdots,n\}$。以前所说矩阵的初等变换：倍法变换、消法变换、换法变换，按这里关于变换的定义，都是集合$M_{m,n}(F)$的变换，而且都是一一变换。而且不难指出这些一一变换各自的逆变换。我们回忆一下以前关于初等变换的符号。我们曾经用符号 $D_i(\lambda),\lambda\neq 0;P_{ij}(k)$；$C(i,j)$分别表示倍法变换、消法变换和换法变换。于是知道：

$D_i(\lambda)$的逆变换是$D_i(\lambda^{-1})$；

$P_{ij}(k)$的逆变换是$P_{ij}(-k)$；

$C(i,j)$的逆变换是$C(i,j)$。

特别地，当$m=n$ 时，对n阶方阵的集合 $M_n(F)$ 来说，熟知的规则

$$\varphi:(a_{ij})\longmapsto(a_{ij})',$$

即规则φ把n阶方阵(aij)射成它的转置阵 $(a_{ij})'$，是$M_n(F)$的一个变换，而且是一一变换。显然φ的逆变换就是φ自己。

例 8　考虑实系数多项式的集合，即集合 $D[x]=\{f(x)\mid f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n,\ a_i\in D\}$。规则

$$\varphi:f(x)\longmapsto f'(x),$$

即规则φ把多项式 $f(x)$射成它的导式 $f'(x)$。显然，φ 是$D[x]$的一个变换，而且是满变换，但不是单变换。

最后用下边的例子结束这一节的讨论。

例 9　考虑$D[x]$与D。我们给出一个$D[x]$到D的规则

$$\varphi:f(x)\longmapsto f(x)\text{的实根}.$$

这个规则φ就是把多项式 $f(x)$ 射成为它的实根。我们容易指出，这个规则φ不能成为$D[x]$到D的映射。因为 $f(x)\in D[x]$，即 $f(x)$ 是实系数多项式，而实系数多项式未必有实根。于是对于没有实根

的多项式 $f(x)$，如 $f(x)=x^2+1$，规则φ就没能在D中给 $f(x)$ 规定出象来。因此，按规则φ，不能对$D[x]$的每一个元素在D中确定出唯一的一个元素来，所以φ不是$D[x]$到D的映射。此外，从另一个侧面也能指出φ不是$D[x]$到D的映射。即对于 $D[x]$ 中有实根的多项式 $f(x)$，由于 $f(x)$的实根可能不止一个，如 $f(x)=x^2-3x+2$就有两个实根 1 与 2，于是这个规则φ说不清楚它把这样的 $f(x)$给射成为哪一个实根。因此，按规则 φ，对实根个数多于一个的多项式 $f(x)$在D中确定出的元素不是唯一的，所以φ不是 $D[x]$ 到D的映射。

虽然如此，以上讨论倒不是毫无启发的。事实上，规则φ之所以不能成为映射，不外出自两个方面的漏洞。一是对于没有实根的多项式，φ没能给规定出象来，再就是对于有不同实根的多项式，φ给规定的象不是唯一的。针对这种情况，如果我们把 $D[x]$ 中每一个多项式 $f(x)$ 的全部实根做为一个整体来考虑，那么它就是一个由 $f(x)$ 所唯一确定的实数集——D的子集。这时就连没有实根的多项式也不例外，只不过其实根集合是一个空集就是了。这样，如果把 $f(x)$与其实根的集合联系起来，即给出$D[x]$到D的幂集$P(D)$一个规则，不妨还用φ来表示：

$$\varphi: f(x) \longmapsto R_f=\{a \mid a \in D,\ f(a)=0\}.$$

于是这个规则φ就是$D[x]$到$P(D)$的一个映射。容易知道，映射φ既不是单的也不是满的。

练 习 二

1. 设C为复数集。举出C的如下的变换各一个：

1）单变换但不是满变换；

2）满变换但不是单变换；

3）既不是单变换也不是满变换；

4）单满变换。

2. 设D为实数集，D^+表示正实数集．举出D到D^+的一个映射，使其有逆映射并求出这个逆映射．

3. 举出一个Z到N的映射，使其有逆映射并求出这个逆映射．

4. 考虑一元多项式的集合$F[x]$．规则

$\varphi_1: f(x) \longmapsto f(x+h)$，$h$为$F$中的定数；

$\varphi_2: f(x) \longmapsto f(kx)$，$k$为$F$中不等于零的数．

说明φ_1, φ_2都是一一变换．

§3 代数运算

数学中，运算是个很基本、很普遍的概念和方法，也是我们比较熟悉的，如数的四则运算．在前面几章里，我们还讨论过多项式的运算、n 维向量的运算．在后面几章里，将会对许多其它的对象讨论运算的问题．正如前面我们说过的那样，把运算做为讨论的基本课题是高等代数的一个显著特点．因此有必要把有关运算的一些概念明确起来．这一节就是用映射的概念来说明什么叫运算．

回顾一下，在代数学的范围里，我们讨论过的各式各样的运算——数的加、减、乘、除法；多项式的加、减、乘法；向量的加法，数与向量的乘法等等．所有这些运算有没有什么共同之处，它们的共同本质又是什么呢？很简单，所有这些运算无一例外，它们的最基本的特征都是："两得一个"，事实上，这也就是代数运算这一重要概念的最根本的东西．

设A，B，C是任意三个非空集合，它们有的可以是相同的集合．

定义 **如果φ是积集$A \times B$到C的一个映射，则称φ是A与B到C的一个（代数）运算．特别地，当$C = B = A$时，就把A与A到A的运算简称为集合A的运算．**

如果φ是A与B到C的一个代数运算，按定义，这φ就是积集

$A\times B$到C的一个映射．因此，对于每一个元素 $(a,b)\in A\times B$，通过映射φ，都射成C中一个唯一确定的元素c，即

$$\varphi: (a,b)\longmapsto c, \text{或者 } \varphi((a,b))=c.$$

按照通常运算符号的用法，我们把(a,b)在φ之下的象等于c 这一事实改成以下写法：

$$a\varphi b=c,\ \forall\ a\in A,\ b\in B,\ c\in C$$

例 1　设φ是实数集D的一个运算，其运算规则为

$$\varphi: (a,b)\longmapsto c,\ \text{而} c=a+b.$$

这个运算规则φ就是通常的加法，即

$$c=a\varphi b=a+b.$$

例 2　令$\dot{D}=D-\{0\}$，即$\dot{D}$是非零实数集．考虑D与$\dot{D}$到D的一个运算φ，其运算规则为

$$\varphi: (a,b)\longmapsto c=\frac{a}{b},\ \text{此处 } a\in D,\ b\in\dot{D}.$$

这个运算规则 φ 就是通常的除法，即对任意的实数 $a,b\neq 0$，都有 $a\varphi b=\frac{a}{b}$.

例 3　考虑实数集D上n维向量的集合， 即 $D^{(n)}=\{(x_1,x_2,\cdots,x_n)\mid x_i\in D\}$．令$\varphi$是$D$与$D^{(n)}$到$D^{(n)}$的一个运算，其运算规则为

$$\varphi: (k,(x_1,x_2,\cdots,x_n))\longmapsto(kx_1,kx_2,\cdots,kx_n).$$

这个运算规则φ就是数乘向量的倍数运算，即

$$k\varphi(x_1,x_2,\cdots,x_n)=k\cdot(x_1,x_2,\cdots,x_n)=(kx_1,kx_2,\cdots,kx_n).$$

以上各例主要说明，以往我们所熟悉的一些运算，都是这里所说的代数运算的具体实例，下面我们再以代数运算的概念来观察其它一些与代数运算有关的数学现象．

例 4　考虑自然数集合

$$N: 1,\ 2,\ 3,\ \cdots,\ n,\cdots.$$

令φ_1是N的一个代数运算，其运算规则为

$$a\varphi b=c=(a,b),$$

其中 $a, b, c \in N$，(a,b) 是a与b的最大公约数．这就是说，求两个正整数a与b的大公约的方法是正整数集N的一个代数运算．其实这并不费解，而是很自然的事情．因为求a与b的大公约，和求a与b的和 $a+b$、a与b的积$a\cdot b$设什么两样，正是"两得一个"的一种方法．同样地，求两个多项式 $f(x)$ 与$g(x)$的最高公因式$(f(x),g(x))$的方法也可以看做是多项式集合$F[x]$的一个代数运算．

例 5　设A为任一集合，考虑A的幂集，即以A的子集为元素组成的集合$P(A)=\{S|\ S\subseteq A\}$．令$\varphi$是$P(A)$的一个代数运算，其运算规则为

$$\varphi:(S_1,\ S_2)\longmapsto S_3=S_1\cup S_2,\quad 或S_1\varphi S_2=S_1\cup S_2,$$

即S_1与S_2在φ之下的运算结果为其并集$S_1\cup S_2$．这就是说，求并集的方法是幂集$P(A)$的代数运算．

类似地求交集的方法也是幂集 $P(A)$的代数运算．

例 6　考虑实数域D上的三维向量组成的集合 $D^{(3)}=\{(x_1,x_2,x_3)|\ x_1,\ x_2,\ x_3\in D\}$．我们给出 $D^{(3)}$与 $D^{(3)}$ 到D的代数运算φ，其运算规则为

$$\varphi:((a_1,a_2,a_3),\ (b_1,b_2,b_3))\longmapsto a_1b_1+a_2b_2+a_3b_3,$$

或者写成

$$(a_1,a_2,a_3)\varphi(b_1,b_2,b_3)=a_1b_1+a_2b_2+a_3b_3.$$

这个运算规则就是，两个三维向量$(a_1,\ a_2,\ a_3)$与$(b_1,\ b_2,\ b_3)$在φ之下的运算结果为它们的对应分量乘积之和．显然，两个向量的运算结果是一个数．这个运算规则在解析几何中有重要的作用．通常把这个代数运算叫做数积运算或内积运算．

例 7　设 $A=\{d\}$，即A是只含一个元素d的集合．再考虑实系数多项式的集合$D[x]$．

我们给出一个A与$D[x]$到 $D[x]$ 的代数运算φ，其运算规则为

$$\varphi:(d,f(x))\longmapsto f'(x),$$

或者写成

$$d\varphi f(x)=f'(x).$$

这个运算规则就是，d 与任一多项式 $f(x)$ 的运算结果为 $f(x)$ 的导式 $f'(x)$。

这个例子的主要特点在于集合A只有一个元素，记作d。因此，所谓的A与$D[x]$到$D[x]$的运算，其实就是这一个元素 d 和任意一个多项式 $f(x)$ 进行运算。这样，如果我们把体现运算规则的记号 φ 略去不写，那么就有

$$d\varphi f(x)=df(x)=f'(x).$$

于是很容易和我们非常熟悉的一种记号联系起来，即

$$\frac{d}{dx}f(x)=f'(x).$$

综上所述，此例中的代数运算 φ 就是求导数的方法——微分运算。

通过以上诸例，更加充分地看到代数运算"两得一个"的基本特征。并且我们可以这样说：对集合A给定的一种规则（方法）φ，使得A中任意两个有次序的元素a与b，按照给定的规则 φ，能在A中确定一个唯一的元素c，记作$a\varphi b=c$，那么φ就是集合A的一个代数运算。由此，在我们考虑对某一集合建立代数运算时，只须给出一个"两得一个"的方法就行了，不必从$A\times A$ 到A的映射说起。

按代数运算的定义和以上各例，可以看出，给一个集合建立代数运算是有很大的随意性的。换句话说，要想给一个集合建立代数运算并不难，随便就可办到。但是可以理解，随便建立起来的运算未必有多大用处。实际上，数学中最基本的运算——数的四则运算，都是满足一定的运算规律的。而一般集合的运算也只有具备某种规律时才有加以研究的必要。对于一个代数运算所能具有的规律是各式各样的，其中最基本最常见的就是我们所熟悉的三条运算规律：结合律、交换律、分配律。下面我们对一般的代数运算来说明这三条算律的意义。

设A为任一非空集合，我们用一个小圈"○"表示A的任意一个代数运算。当然，在A是某一具体的集合，所说的运算又是大家已知的一个具体的运算时，这时符号"○"就自然应该具体化为

通常所惯用的运算符号．比如例1里的运算φ就是通常的加法，因此这个运算符号φ就应被惯用的加法符号“+”所代替．再如例2里的运算φ就是通常的除法，因此这个运算符号φ就自然应该改写成惯用的除法符号“÷”，即应把$a\varphi b$写成$a\div b$．

如果对A中任意三个元素a，b，c等式

$$(a\bigcirc b)\bigcirc c=a\bigcirc(b\bigcirc c)$$

都成立，则称运算“○”满足（适合）结合律．

如果对A中任意两个元素a,b等式

$$a\bigcirc b=b\bigcirc a$$

都成立，则称运算“○”满足（适合）交换律．

设$\oplus$与$\otimes$是集合A的两个运算，如果对A中任意三个元素a,b,c等式

$$a\otimes(b\oplus c)=(a\otimes b)\oplus(a\otimes c),$$

$$(b\oplus c)\otimes a=(b\otimes a)\oplus(c\otimes a)$$

都成立，则称运算$\otimes$对运算$\oplus$满足（适合）分配律．当只有前（后）一个等式成立时，就说$\otimes$对$\oplus$满足（适合）左（右）分配律．

例8　考虑整数集Z．数的加法、减法、乘法都是Z的代数运算．我们早已熟知，Z的加法与乘法这两个运算既满足结合律又满足交换律．但是，整数的减法却刚好相反：既不满足结合律又不满足交换律．事实上对于3，2，1这三个整数来说，就有

$$(3-2)-1=1-1=0,$$

$$3-(2-1)=3-1=2,$$

即$(3-2)-1\neq 3-(2-1)$，所以整数减法不满足结合律．再由$3-2=1$，而$2-3=-1$，因此$3-2\neq 2-3$．这又说明整数减法也不满足交换律．

同样早已知道的，乘法对加法满足分配律；乘法对减法也满足分配律．

例9　N为正整数集．用“○”表示N的求最大公约数的运算，即

$$a \bigcirc b=(a,b), \forall a,b \in N.$$

于是

$$(a \bigcirc b) \bigcirc c=a \bigcirc (b \bigcirc c),$$

$$a \bigcirc b=b \bigcirc a,$$

即求大公约的运算既满足结合律又满足交换律。后者是自明的，因为a，b的大公约与这两个数的次序没有关系。前者由最大公约的特性：a与b的公约数d是最大的当且仅当a与b的任意一个公约数u都是d的一个约数，也不难得到说明。

同样可知，求最小公倍数的方法是集合N的一个代数运算，并且既满足结合律又满足交换律。

例10 由例5已知，对任一集A的幂集 $P(A)$来说，求并的方法"∪"求交的方法"∩"都是 $P(A)$的代数运算。这里我们指出，这两个运算都满足结合律，也都满足交换律。并且"∪"对"∩"满足分配律，"∩"对"∪"也满足分配律。即有下列等式成立：

$$(S_1 \cup S_2) \cup S_3=S_1 \cup (S_2 \cup S_3), \quad S_1 \cup S_2=S_2 \cup S_1;$$

$$(S_1 \cap S_2) \cap S_3=S_1 \cap (S_2 \cap S_3), \quad S_1 \cap S_2=S_2 \cap S_1;$$

$$S_1 \cup (S_2 \cap S_3)=(S_1 \cup S_2) \cap (S_1 \cup S_3),$$

$$S_1 \cap (S_2 \cup S_3)=(S_1 \cap S_2) \cup (S_1 \cap S_3), \quad S_1, \ S_2, \ S_3 \in P(A).$$

其中最后一个等式已于§1的例5证明过了，另外的五个等式也都容易类似地得到验证。

有一点应当指出的，即"∪"对"∩"及"∩"对"∪"都满足分配律。这一点对于数的加法、乘法来说就不这样，即加法对乘法不满足分配律。事实上对3，2，1这三个数来看就有

$$3+(2 \times 1) \neq (3+2) \times (3+1).$$

本节最后我们讨论结合律、交换律、分配律的一般意义。

首先考虑结合律。

设A为任一非空集合。用"◯"表示A的一个给定的代数运算。

我们知道，代数运算"◯"能使A中元素"两得一个"。由此

不难办到，对A中任意n个元素($n>2$) 都能算得一个元素。例如，对A中三个元素 a，b，c 来说，就可用两种方式算出结果来：

$$(a\bigcirc b)\bigcirc c \text{ 或者 } a\bigcirc(b\bigcirc c).$$

对A中四个元素 a，b、c，d 来说（在a，b，c、d这样的次序下）算出一个结果的方式就更多些，实际上有以下五种方式：

$$((a\bigcirc b)\bigcirc c)\bigcirc d,\ (a\bigcirc b)\bigcirc(c\bigcirc d),\ (a\bigcirc(b\bigcirc c))\bigcirc d,$$
$$a\bigcirc((b\bigcirc c)\bigcirc d),\ a\bigcirc(b\bigcirc(c\bigcirc d)).$$

这样就出现一个问题：对同样一串元素，由于进行运算的具体方式不同，算得的结果会怎样呢？当然，如果运算"○"满足结合律，那么对三个元素 a，b，c 来说用可能有的两种不同方式算得的结果保证是一样的。可是对于四个元素的情形，由五种可能的具体方式算得的结果也准都相同吗？当元素的个数更多时，能够算出结果的方式也就更多些，此时算得的众多结果又如何？从字面上看结合律并没回答这个问题。而实际上，结合律已经解决了这个问题。这就是下面要证明的结论。

为了叙述方便，我们把上述用逐步添加括号的方式，从若干个元素算得一个元素的过程简称为"加括号步骤"。

命题 1　设集合A的代数运算"○"满足结合律，那么对A中任意n个元素 $a_1,a_2,\cdots,a_n$，在这个次序下用任何一种加括号步骤算得的结果都一样。

证明　对元素的个数n用第二数学归纳法。

$n=3$时命题本身就是结合律，因此命题自然成立。

假设对于元素个数 $<n$ 的情形命题成立，去证对于n个元素的情形命题也成立。

考虑 a_1，$a_2,\cdots,a_n$的所有可能的加括号步骤。其中有一种最简单的就是从左逐次往右算出结果来，即

$$(\cdots((a_1\bigcirc a_2)\bigcirc a_3)\bigcirc\cdots)\bigcirc a_n.$$

我们用 $\pi(a_1\bigcirc\cdots\bigcirc a_n)$ 表示用任何一种加括号步骤从 a_1，a_2，$\cdots,a_n$ 算得的结果。于是我们去证

$$\pi(a_1\bigcirc\cdots\bigcirc a_n)=(\cdots((a_1\bigcirc a_2)\bigcirc a_3)\bigcirc\cdots)\bigcirc a_n.$$

对于 $\pi(a_1\bigcirc\cdots\bigcirc a_n)$ 来说，不论是怎样一种加括号步骤，其最后一步都是从两个元素算出结果的. 即

$$\pi(a_1\bigcirc\cdots\bigcirc a_n)=b_1\bigcirc b_2,$$

这里的b_1是前 $i(i<n)$ 个元素 a_1, $\cdots$, a_i用某种加括号步骤算得的结果，b_2是后 $n-i$ 个元素$a_{i+1},\cdots,a_n$用某种加括号步骤算得的结果. 于是

$$\begin{aligned}\pi(a_1\bigcirc\cdots\bigcirc a_n)&=\pi(a_1\bigcirc\cdots\bigcirc a_i)\bigcirc\pi(a_{i+1}\bigcirc\cdots\bigcirc a_n)\\&=\pi(a_1\bigcirc\cdots\bigcirc a_i)\bigcirc(\pi(a_{i+1}\bigcirc\cdots\bigcirc a_{n-1})\bigcirc a_n)\\&=(\pi(a_1\bigcirc\cdots\bigcirc a_i)\bigcirc\pi(a_{i+1}\bigcirc\cdots\bigcirc a_{n-1}))\bigcirc a_n\\&=\pi(a_1\bigcirc\cdots\bigcirc a_{n-1})\bigcirc a_n\\&=((\cdots((a_1\bigcirc a_2)\bigcirc a_3)\bigcirc\cdots)\bigcirc a_{n-1})\bigcirc a_n.\end{aligned}$$

这样就证明了命题 1.

我们把从 a_1, a_2, $\cdots$, a_n用任意加括号步骤算得的唯一结果简记作 $a_1\bigcirc\cdots\bigcirc a_n$.

按命题 1，只要运算满足结合律，那么对于任意n个元素 a_1, a_2, $\cdots,a_n$, 在这个次序下，用随便一种加括号步骤都能算出相同的结果，这在做具体计算时无疑是很大的方便. 结合律的意义就在于此.

用加括号步骤做运算时，我们总是强调“在这样的次序下”的限制. 这是因为改变元素次序对运算结果一般是有影响的. 但是，如果运算满足交换律，这样的限制就完全可以取消. 这就是下面要证明的结果.

命题 2　如果A的运算“$\bigcirc$”满足结合律又满足交换律，那么对A中的元素 a_1, $a_2,\cdots$, a_n, 在任意次序下用加括号步骤算得的结果都一样.

证明　对n做数学归纳法.

$n=2$时，命题自然成立.

假设元素个数 $<n$ 时命题成立，去证n个元素的情形命题也成

立。

设$i_1, i_2, \cdots, i_n$是 1，2，…，n 的任一排列。我们证明：

$$a_{i_1}\bigcirc a_{i_2}\bigcirc\cdots\bigcirc a_{i_n}=a_1\bigcirc a_2\bigcirc\cdots\bigcirc a_n.$$

在 $i_1, i_2, \cdots, i_n$中有一个是n，不妨令 $i_k=n$.

$$\begin{aligned}a_{i_1}\bigcirc\cdots\bigcirc a_{i_k}\bigcirc\cdots\bigcirc a_{i_n}&=a_{i_1}\bigcirc\cdots\bigcirc a_n\bigcirc\cdots\bigcirc a_{i_n}\\&=(a_{i_1}\bigcirc\cdots\bigcirc a_{i_{k-1}})\bigcirc[a_n\bigcirc(a_{i_{k+1}}\bigcirc\cdots\bigcirc a_{i_n})]\\&=(a_{i_1}\bigcirc\cdots\bigcirc a_{i_{k-1}})\bigcirc[(a_{i_{k+1}}\bigcirc\cdots\bigcirc a_{i_n})\bigcirc a_n]\\&=[(a_{i_1}\bigcirc\cdots\bigcirc a_{i_{k-1}})\bigcirc(a_{i_{k+1}}\bigcirc\cdots\bigcirc a_{i_n})]\bigcirc a_n\\&=(a_1\bigcirc\cdots\bigcirc a_{n-1})\bigcirc a_n\\&=a_1\bigcirc\cdots\bigcirc a_{n-1}\bigcirc a_n.\end{aligned}$$

这就证明了命题 2 .

交换律的意义是清楚的，特别是与结合律一起，它的作用就更明显了。以后我们会遇到不满足交换律的运算，那时将会感到计算起来很不方便，这可以说明交换律也是一条很重要的算律。

最后证明关于分配律的

命题 3　设$\odot$，$\oplus$是A的两个运算，$\oplus$满足结合律，$\odot$对$\oplus$满足分配律。那么

$$a\odot(b_1\oplus\cdots\oplus b_n)=(a\odot b_1)\oplus\cdots\oplus(a\odot b_n),$$
$$(b_1\oplus\cdots\oplus b_n)\odot a=(b_1\odot a)\oplus\cdots\oplus(b_n\odot a).$$

其中a，$b_1, \cdots, b_n$ 为A中任意元素。

证明　对n用数学归纳法。

$n=2$时，命题自然成立。

假设对于元素b_i的个数$<n$时命题成立，去证b_i 的个数为n时命题也成立。于是

$$\begin{aligned}a\odot(b_1\oplus\cdots\oplus b_n)&=a\odot[(b_1\oplus\cdots\oplus b_{n-1})\oplus b_n]\\&=[a\odot(b_1\oplus\cdots\oplus b_{n-1})]\oplus(a\odot b_n)\\&=[(a\odot b_1)\oplus\cdots\oplus(a\odot b_{n-1})]\oplus(a\odot b_n)\\&=(a\odot b_1)\oplus\cdots\oplus(a\odot b_{n-1})\oplus(a\odot b_n).\end{aligned}$$

同理可证另一个等式成立。这样就证明了命题 3 .

由命题 3，不难说明下列等式成立：

$$(a_1 \oplus \cdots \oplus a_m) \odot (b_1 \oplus \cdots \oplus b_n)$$
$$= (a_1 \odot b_1) \oplus \cdots \oplus (a_1 \odot b_n) \oplus \cdots \oplus (a_m \odot b_1) \oplus \cdots \oplus (a_m \odot b_n)$$
$$= (a_1 \odot b_1) \oplus \cdots \oplus (a_m \odot b_1) \oplus \cdots \oplus (a_1 \odot b_n) \oplus \cdots \oplus (a_m \odot b_n).$$

这就是通常展开括号的方法。

分配律的意义在于它把一个集合中的两个运算连系起来，通过这种连系能在运算过程中改变两个运算的次序。

练 习 三

1．设 $A=\{-1,1\}$．在A中建立两个代数运算．你给出的这两个运算满足哪些算律？

2．设 $A=\{0,1,2\}$．在A中建立两个代数运算，你给出的这两个运算满足哪些算律？

3．试在二阶方阵的集合$M_2(F)$里建立两个代数运算．你给出的这两个运算满足哪些算律？

4．关于自然数集N的规则

$$\varphi: (a,b) \longmapsto a^b$$

是N的一个代数运算．它是否满足结合律，交换律？

5．设D^+为正实数集．令规则

$$\varphi_1: (a,b) \longmapsto \frac{a+b}{2},$$

$$\varphi_2: (a,b) \longmapsto (ab)^{\frac{1}{2}},$$

那么 φ_1，φ_2都是D^+的代数运算．φ_1，φ_2是否满足结合律、交换律？

§4 集合按子集的分类

我们已经知道，映射是两个集合之间建立联系的一种方法．利用这种联系来对两个集合进行比较，通过这种比较可以做到由一个

集合的性质去推测另一个集合可能有的性质。除了这种通过相互之间的联系，由此及彼地认识事物的方法之外，有时也要把一个集合分成若干个子集来加以研究，通过子集的性质来掌握整个集合的性质。这种由部分到整体、分门别类的来认识事物也是研究问题的一种重要方法。下面我们讲两个问题：什么叫集合的分类（分解），怎样对一个集合做分类。

先看两个例子：

例1 考虑整数集 $Z=\{\cdots,-4,-3,-2,-1,0,1,2,3,4\cdots\}$。取定一个正整数 $m=4$。于是我们可以如下的得到 Z 的四个子集：

$$r_0=\{n\in Z \mid n=4q\},$$

$$r_1=\{n\in Z \mid n=4q+1\},$$

$$r_2=\{n\in Z \mid n=4q+2\},$$

$$r_3=\{n\in Z \mid n=4q+3\}.$$

Z 的这四个子集具有以下重要性质：

1） $r_i\neq\phi$，$i=0,1,2,3$；

2） $r_i\cap r_j=\phi$，$i\neq j$；

3） $Z=r_0\cup r_1\cup r_2\cup r_3$。

这说明整数集 Z 恰好分成四个两两不相交的非空子集的并。一般地，对任一给定的正整数 m，整数集 Z 都能按上述方法分成 m 个两两不相交的非空子集的并，其中每一子集恰好是以 m 去除余数相同的一切整数组成的。

例2 考虑数域F上二阶方阵做成的集合 $M_2(F)=\{(a_{ij})\mid a_{ij}\in F;\ i,j=1,2\}$。我们按以下方法确定 $M_2(F)$ 的三个子集：

$$r_0=\{(a_{ij})\in M_2(F)\mid \operatorname{rank}(a_{ij})=0\},$$

$$r_1=\{(a_{ij})\in M_2(F)\mid \operatorname{rank}(a_{ij})=1\},$$

$$r_2=\{(a_{ij})\in M_2(F)\mid \operatorname{rank}(a_{ij})=2\}.$$

$M_2(F)$ 的这三个子集具有以下重要性质：

1） $r_i\neq\phi$，$i=0,1,2$；

2） $r_i\cap r_j=\phi$，$i\neq j$；

3） $M_2(F)=r_0\cup r_1\cup r_2$.

这说明二阶方阵的集合$M_2(F)$恰好分成三个两两不交的非空子集的并，这里的每一个子集恰好是由秩数相同的二阶方阵组成的。

一般地我们给出以下的

定义1 设A为任一集合，$\Sigma=\{A_1,A_2,\cdots\}$是$A$的一组子集。如果满足

1） $A_i\neq\phi$，$A_i\in\Sigma$；

2） $A_i\cup A_j=\phi$，$i\neq j$；

3） $A=A_1\cap A_2\cup\cdots$，

则称$\Sigma=\{A_1,A_2,\cdots\}$是$A$的**一个分类（分解）**，每一个子集$A_i$叫做在分类$\Sigma$之下的**一个类**。

按此定义，例1中整数集Z的四个子集组成的 $\Sigma=\{r_0,r_1,r_2,r_3\}$是Z的一个分类。于是 r_0，r_1，r_2，r_3是在这个分类之下的四个类。在例2中，$M_2(F)$的三个子集组成的$\Sigma=\{r_0,r_1,r_2\}$就是$M_2(F)$的一个分类，$\Sigma=\{r_0, r_1, r_2\}$把集合 $M_2(F)$ 分成为三个类：r_0，r_1，r_2。

显然可以这样理解：所谓集合A的一个分类Σ，就是说Σ是幂集 $P(A)$ 的一个子集：$\Sigma\subseteq P(A)$，并且Σ中的元素满足上述的三个条件。

下面再举几个例子。

例3 对 $M_2(F)$ 考虑它的以下两个子集：

$$d_0=\{(a_{ij})\in M_2(F)\ |\det(a_{ij})=0\},$$

$$\bar{d}_0=\{(a_{ij})\in M_2(F)\ |\det(a_{ij})\neq 0\}.$$

容易看出$\Sigma_1=\{d_0,\bar{d}_0\}$ 是$M_2(F)$的一个分类，即把全体二阶方阵分成两类：行列式等于0的是一类，行列式不等于0的是一类。

显然，这个分类比较粗略。如果我们把行列式不为0的二阶方阵再进行细分，即令

$$d_a=\{(a_{ij})\in M_2(F)\ |\det(a_{ij})=a\}.$$

那么$\Sigma_2=\{d_a\ |a\in F\}$也是$M_2(F)$的一个分类。分类$\Sigma_2$就是行列式相

同的分在一类，行列式不相同的不分在同一个类里．看得出来，分类Σ_2比Σ_1精细得多．

例 4　对F上的一元多项式的集合$F[x]$考虑它的以下一些子集：

$$d_i=\{f(x)\in F[x]|\deg f(x)=i\},\ i=0,\ 1,\ 2,\cdots.$$

这里子集d_i就是次数等于i的多项式组成的子集．显然，这些子集满足分类定义的前两个条件，即每一个子集d_i都不空；每两个不同的子集d_i与d_j（$i\neq j$）都不相交，但是不满足这个定义的第三个条件，即这些子集的并集$\bigcup\limits_{i=0,1,2\cdots}d_i$不等于$F[x]$．因为零多项式没有次数，所以零多项式不属于任何一个子集d_i．这样，$\{d_0,\ d_1,\ d_2,\ \cdots,\}$不能构成$F[x]$的一个分类，但是，只要把零多项式做为一个子集$\{O\}$补充进去，即令

$$\Sigma=\{\{0\},\ d_0,d_1,d_2,\cdots\},$$

那么Σ就是$F[x]$的一个分类．

例 5　设$A=\{ax^2+bx+c|\ a,b,c\in D,a\neq 0\}$．即$A$是一切实系数一元二次多项式组成的．考虑$A$的下列三个子集：

$$r_0=\{ax^2+bx+c\in A|\ b^2-4ac=0\},$$
$$r_1=\{ax^2+bx+c\in A\ |b^2-4ac>0\},$$
$$r_2=\{ax^2+bx+c\in A|\ b^2-4ac<0\}.$$

这里r_0是有等根的二次多项式组成的；r_1是有相异二实根的二次多项式组成的；r_2是没有实根的二次多项式组成的．这样，$\Sigma=\{r_0,\ r_1,r_2\}$构成A的一个分类．

从以上各例可以看出，对集A的每一个确定的分类Σ来说，凡是同在一类里的元素都具有一种共同的性质，而在不同两类里的元素所具有的性质也不同．比如，例 1 中的分类Σ，同在一类的整数，以 4 除之所得余数都相同；而在不同两类的整数，以 4 除之所得余数则不同．又如例 2 中的分类Σ，同在一类的二阶方阵秩数都相同，而在不同两类的二阶方阵秩数则不同．

对于集合的分类所呈现的这种规律性也可以从另外一种角度来揭示。即凡同在一类里的元素都有某种关系；而不在同一类里的元素则没有这种关系。比如，例 1 中的分类Σ，同在一类里的任二整数之差都是 4 的倍数；而不在同一类里的任二整数之差则不是 4 的倍数。又如，例 2 中的分类 Σ，同在一类里的任意两个二阶方阵都相抵；而不在同一类里的任意两个二阶方阵则不相抵。一般来说，集合的任何一种分类都是利用元素间的某种关系得到的。这就是我们要讲的第二个问题。

设 A 为任一非空集合，符号"～"表示 A 中元素 a 与 b 之间的一种关系。若 a 与 b 有这种关系，就记作 $a \sim b$；若 a 与 b 没有这种关系，就记作 $a \nsim b$。

设"～"是集 A 的元素之间的一种关系。如果对任意的 $a,b \in A$，$a \sim b$ 与 $a \nsim b$ 有且只有一种情形成立，则称"～"是集合 A 的一种关系。

例 6 考虑整数集 Z。令～表示两个整数之间的一种关系：如果 a 与 b 的奇偶性相同，则 $a \sim b$；如果 a 与 b 的奇偶性不同，则 $a \nsim b$。比如，$2 \sim 4$，$-1 \sim 1$，$1 \nsim 2$，$3 \sim 3$，$0 \nsim 1$ 等等。因为任二整数的奇偶性或者相同或者不同，二者必居其一且只居其一。这说明整数之间的关系～是整数集 Z 的一种关系。

整数集 Z 的这个关系还可以用以下方式描述得简明些，即

$$\sim:\ a \sim b \text{ 当且仅当 } 2 \backslash a - b.$$

用话来说就是：如果 a 与 b 的差被 2 整除，a 与 b 就有关系，否则 a 与 b 就没有关系。

例 7 考虑 n 阶方阵的集合 $M_n(F)$。令～表示两个 n 阶方阵 A，$B \in M_n(F)$ 之间的一种关系，即

$$\sim:\ A \sim B \text{ 当且仅当 } A \text{ 与 } B \text{ 相抵}.$$

这就是说，如果 A 与 B 相抵，那么 A 与 B 就有关系，否则 A 与 B 就没关系。显然，关系～是集合 $M_n(F)$ 的一种关系。

例 8 考虑一元多项式的集合 $F[x]$。令～是两个多项式之间

的一种关系，即

$$\sim: f(x)\sim g(x) \text{ 当且仅当 } f(x)\mid g(x),\ g(x)\mid f(x).$$

这就是说，如果 $f(x)$ 与 $g(x)$ 互相能整除，那么 $f(x)$ 与 $g(x)$ 就有关系，否则就没有关系。显然，关系～是 $F[x]$ 的一种关系。

例 9　考虑有理数集合 Q。令～是两个有理数之间的一种关系。即

$$\sim: \frac{a}{b}\sim\frac{c}{d} \text{ 当且仅当 } ad=bc。$$

这就是说，如果 $\frac{a}{b}$ 与 $\frac{c}{d}$ 是相等的有理数，那么 $\frac{a}{b}$ 与 $\frac{c}{d}$ 就有关系，否则就没有关系。显然，关系～是有理数集的一种关系。

例10　对整数集 Z 考虑两个整数之间如下的一些关系：

$\sim_1$：$a\sim_1 b$ 当且仅当 $a\mid b$；

$\sim_2$：$a\sim_2 b$ 当且仅当 $a<b$；

$\sim_3$：$a\sim_3 b$ 当且仅当 $ab>0$；

$\sim_4$：$a\sim_4 b$ 当且仅当 $ab\leqslant 0$；

$\sim_5$：a 与 b 的符号如果相同，则规定 a 与 b 有关系，即 $a\sim_5 b$；如果 a 与 b 的符号相反，则规定 a 与 b 没有关系，即 $a\nsim_5 b$。

这里，对关系 $\sim_1$ 来说就是，如果 $a\mid b$ 那么 $a\sim_1 b$，否则 $a\nsim_1 b$。比如 $2\sim_1 4$，$2\nsim_1 3$，$4\nsim_1 2$，$-1\sim_1 -1$ 等等。因为任二整数 a 与 b 或者 $a\mid b$ 或者 $a\nmid b$，并且只有一种可能，从而 $a\sim_1 b$ 与 $a\nsim_1 b$ 恰有一种情形成立，所以关系 $\sim_1$ 是整数集 Z 的一种关系。同样可以说明关系 $\sim_2$、$\sim_3$ 和 $\sim_4$ 各是整数集 Z 的一种关系。下面我们考察一下关系 $\sim_5$。这个关系是利用整数的正负性来规定的。具体的说就是：如果 a 与 b 都是正的或者都是负的，那么 a 与 b 就有关系：$a\sim_5 b$；如果 a 与 b 中有一个是正的而另一个是负的，那么 a 与 b 就没有关系：$a\nsim_5 b$。比如 $3\sim_5 1$，$2\sim_5 1$，$1\sim_5 1$，$-1\sim_5 -1$，$-1\nsim_5 1$，$-6\nsim_5 6$ 等等。进而我们自然要问：0 与 1，在关系 $\sim_5$ 之下，是有关系还是没有关系？由于关系 $\sim_5$ 是利用整数的正负性给两个整数规定了是有关系或者是没有关系。然

而整数0既不是正的也不是负的。因此，关系 $\sim_5$ 对于这个没有正负性可言的整数0没给规定与任何一个整数是有关系或者没有关系。这样，关系 $\sim_5$ 就不是整数集Z的一种关系。建议读者把这个关系 $\sim_5$ 与关系 $\sim_3$ 比较一下，从中会得到一些启发的。

能够给集合构成分类的是一种特殊的关系。下面我们就来介绍这种关系。

定义2 设～是集合A的一种关系。如果以下条件成立：

1）**反身性** 对任一 $a\in A$，都有 $a\sim a$；

2）**对称性** 若$a\sim b$则$b\sim a$，

3）**传递性** 若$a\sim b$，$b\sim c$，则 $a\sim c$，

就说～是集合A的一种等价关系。

容易验证上述例6—例9中的关系都是等价关系。例10中的五个关系都不是等价关系。

事实上 关系 $\sim_1$ 不具有对称性； 关系 $\sim_2$ 不具有反身性也不满足对称性；关系 $\sim_3$ 不具有反身性（为什么？）；关系 $\sim_4$ 不满足传递性；关系 $\sim_5$ 根本就不是Z的关系，自然更谈不上是不是等价关系。

下面证明本节的主要结果。

定理 集合A的一个分类决定A的一个等价关系。反过来，集合A的一个等价关系决定A的一个分类。

证明 先证第一个结论。设Σ是A的任意一个分类。我们利用A的这个分类Σ，给出A的元素之间的一个关系，记作“～”，即

$$a\sim b \text{当且仅当} a \text{与} b \text{在同一类里。}$$

换句话说就是：如果a与b被分类Σ分在一个类里，那么就规定a与b有关系：$a\sim b$；如果a与b被Σ分在不同的类里，那么就规定a与b没有关系：$a\not\sim b$。

因为对A的任一分类来说，A中每一元素恰好分在一个确定的类里，所以利用分类Σ如上那样规定的元素之间的关系“～”是集合A的一个关系。下面我们指出关系“～”具备等价关系所要求的

三条性质。首先，对任何一个分类来说，自然对每一个元素 a 都有 a 与 a 分在一类，亦即有

1）反身性　对任一 $a\in A$，都有 $a\sim a$。

其次，如果 a 与 b 分在同一类里，当然就有 b 与 a 分在同一类里，亦即

2）对称性　若 $a\sim b$ 则 $b\sim a$。

再次，如果 a 与 b 分在同一类里，b 与 c 分在同一类里，当然就有 a 与 c 分在同一类里，亦即

3）传递性　若 $a\sim b$，$b\sim c$ 则 $a\sim c$。

这样就证明了，利用 A 的分类 Σ 给 A 规定的关系"$\sim$"是 A 的一个等价关系，即 A 的每一个分类都决定 A 的一个等价关系。

再证第二个结论。设"$\sim$"是集合 A 的给定的一个等价关系。我们利用 A 的这个等价关系给 A 造成一个分类。为此，我们首先利用 A 的这个等价关系构造 A 的一些子集。办法如下：

对每一个 $a\in A$，规定 $\bar{a}=\{x\in A|\ x\sim a\}$。

把这样做成的一切不同的子集放在一起，用 Σ 来表示。下面我们指出，Σ 是 A 的一个分类。

首先，因为对任一元素 a 都有 $a\sim a$，所以对我们如上做出的每一个子集 $\bar{a}$ 来说就是 $a\in\bar{a}$，亦即

1）$\bar{a}\neq\phi$，$\forall\bar{a}\in\Sigma$。

其次，Σ 中任二不同子集都不相交。事实上，任取 $\bar{a}$，$\bar{b}\in\Sigma$，且 $\bar{a}\neq\bar{b}$，必有 $\bar{a}\cap\bar{b}=\phi$。假如不是这样，必有 $c\in\bar{a}\cap\bar{b}$。从而 $c\in\bar{a}$ 及 $c\in\bar{b}$。按子集 $\bar{a}$、$\bar{b}$ 的构成定义，则有 $c\sim a$，$c\sim b$。又因关系 $\sim$ 是等价关系，依据对称性，由 $c\sim a$ 可得 $a\sim c$。再依据传递性，由 $a\sim c$，$c\sim b$ 可得 $a\sim b$。这样，对任一 $x\in\bar{a}$，即 $x\sim a$。从而 $x\sim b$，亦即 $x\in\bar{b}$。故有 $\bar{a}\subseteq\bar{b}$。同理可得 $\bar{b}\subseteq\bar{a}$。因而 $\bar{a}=\bar{b}$。这与 $\bar{a}\neq\bar{b}$。相矛盾。所以 $\bar{a}$ 与 $\bar{b}$ 不能含有共公元素，即

2）$\bar{a}\cap\bar{b}=\phi$，$\bar{a}$，$\bar{b}$ 为 Σ 中任二不同元素。

再次，因为对 A 中任一元素 a，都有 $a\in\bar{a}$，所以 A 中每一元素 x

必定属于Σ中某一子集，亦即

$$3)\quad A=\bigcup_{\bar{x}\in\Sigma}\bar{x}.$$

这样就证明了，利用A的已给的等价关系～，如上那样做出的一组子集构成了A的一个分类Σ，即A的一个等价关系决定A的一个分类.

集合的分类与映射有着自然的联系. 因为我们容易指出

命题　集合A的每一个映射决定A的一个分类. A 的每一个分类决定A的一个映射.

证明　假设集合A有一个从A到B的映射φ：$A\xrightarrow{\varphi}B$. 我们利用映射 φ 做出A的一些子集，使其构成A的一个分类. 办法如下：

对任意一个 $a\in A$，规定

$$\bar{a}=\{x\in A|\ \varphi(x)=\varphi(a)\}.$$

换句话说就是，$\bar{a}$是A中与元素a有相同的象的一切元素做成的子集. 把这样做成的一切不同的子集放在一起，用Σ来表示. 下面指出Σ是A的一个分类.

首先，对A中每一元素a，自然有 $\varphi(a)=\varphi(a)$，从而$a\in\bar{a}$. 所以

$$1)\quad \bar{a}\neq\phi,\ \forall\bar{a}\in\Sigma.$$

其次，Σ中任二不同子集都不相交. 不然，令$\bar{a}\neq\bar{b}$，但$\bar{a}\cap\bar{b}\neq\phi$. 于是有$c\in\bar{a}\cap\bar{b}$. 从而 $c\in\bar{a}$，$c\in\bar{b}$. 按 $\bar{a},\bar{b}$的构成定义，则有$\varphi(c)=\varphi(a),\varphi(c)=\varphi(b)$，即有$\varphi(a)=\varphi(b)$. 这就表明：凡与$a$ 有相同的象的元素也与b有相同的象，反之亦然. 亦即

$$\bar{a}\subseteq\bar{b},\ \bar{b}\subseteq\bar{a}$$

从而$\bar{a}=\bar{b}$. 这与$\bar{a}\neq\bar{b}$相矛盾. 所以，对于不同的$\bar{a},\bar{b}$ 不能含有共公元素. 即有

$$2)\quad \bar{a}\cap\bar{b}=\phi,\ \bar{a},\ \bar{b}\text{为}\Sigma\text{中任二不同的元素}.$$

再次，对于A中任一元素a都有$a\in\bar{a}$，所以A中每一元素x必定属于Σ中某一子集，亦即

3) $A=\bigcup\limits_{\bar{x}\in\Sigma}\bar{x}$.

这就证明了，利用A到B的映射φ能够决定A的一个分类.

反之，假设集合A有一个分类Σ，于是这个Σ就是以A的某些子集为元素组成的集合. 因而我们可以考虑从集合A到集合Σ的映射，具体的规则如下：

$$\varphi: a\longmapsto\bar{a},\quad \bar{a}\in\Sigma,\text{ 使 } a\in\bar{a}.$$

换句话说就是，规则φ把A中的元素a射成Σ中含有a的子集$\bar{a}$.

我们下面指出，规则φ是A到Σ的映射. 为此，应当指出两点：A中每个元素都有象；象是唯一的. 这完全由Σ是A的分类决定的，因为A中每一元素a必属于Σ中的一个子集，从而a在φ之下必有象；a不能属于Σ中两个不同的子集，从而a在φ之下的象是唯一的. 这样利用分类Σ规定的方法φ是A到Σ的一个映射，而且是满映射，即$A\xrightarrow{\varphi}\Sigma$.

上述定理和命题，使得关于一个集合A的等价关系、分类、映射三者串连起来，说明三者对于一个集合A起着相同的作用，有同样的意义. 其中分类的观念和方法比较通俗易于理解，因此在以后的讨论中常常是使用分类的语言来说明问题.

前面的例1—例5都是些分类的例子. 因此，依据定理，每一个分类都相应的决定一个等价关系. 比如例1中的分类$\Sigma=\{r_0, r_1, r_2, r_3\}$把整数分成四类. 分类的规则是：$a$与$b$分在一类里当且仅当以4除之，$a$与$b$的余数相同. 于是由这个分类$\Sigma$所决定的Z的一个等价关系“～”，依据定理中的原则规定，应该是

$a\sim b$当且仅当a与b被Σ分在一类里. 换句话说就是，如果a与b分在同一类里，a与b就有关系：$a\sim b$；否则a与b就没有关系：$a\nsim b$.

再如例2中的分类$\Sigma=\{r_0, r_1, r_2\}$把二阶方阵分成三类. 分类

的规则是：二阶方阵(a_{ij})与(b_{ij})分在一类里当且仅当(a_{ij})与(b_{ij})的秩相同．于是由这个分类Σ所决定的$M_2(F)$的一个等价关系"～"，依据定理中的原则规定，应该是

$(a_{ij})\sim(b_{ij})$当且仅当(a_{ij})与(b_{ij})被Σ分在一类，换句话说就是，如果(a_{ij})与(b_{ij})分在同一类里，(a_{ij})与(b_{ij})就有关系：$(a_{ij})\sim(b_{ij})$；否则(a_{ij})与(b_{ij})就没有关系：$(a_{ij})\nsim(b_{ij})$。

例6—例9都是等价关系的例子．因此，依据定理，每一个等价关系都相应的决定一个分类．比如例7中的等价关系～：$(a_{ij})\sim(b_{ij})$当且仅当(a_{ij})与(b_{ij})相抵．于是由这个等价关系决定的$M_n(F)$的一个分类，依据定理中的原则规定，应该是

(a_{ij})与(b_{ij})分在一类里当且仅当$(a_{ij})\sim(b_{ij})$。换句话说就是，如果$(a_{ij})\sim(b_{ij})$，那么(a_{ij})与(b_{ij})就分在同一个类里，否则(a_{ij})与(b_{ij})就不分在同一类里。

练　习　四

1．试给出复数集C的一个分类Σ，使Σ与实数集D之间能建立起一个一一对应．并指出由Σ所决定的等价关系。

2．试给出实数集D的一个分类Σ，使Σ与Z之间能建立起一个一一对应．并指出由Σ所决定的等价关系。

3．把Z做成一个分类Σ，使Σ与偶数集之间能建立起一个一一对应．并指出由这个Σ所决定的等价关系。

4．设$A=\{1,2,3,4,5,6\}$。

1)给A规定一个等价关系$\sim_1$，使由$\sim_1$所决定的分类含有两个子集；

2）给A规定一个等价关系$\sim_2$，使由$\sim_2$所决定的分类含有三个子集；

3）给A规定一个等价关系$\sim_3$，使由$\sim_3$所决定的分类含有一个子集；

4）给A规定一个等价关系 $\sim_4$，使由$\sim_4$所决定的分类含有六个子集。

习　题　七

1．设$A=\{1,2,3,4\}$，$B=\{0,2,4\}$。写出 $A\cup B$，$A\cap B$，$A-B$，$B-A$，$B\times A$，$P(B)$。

2．设$A=(a_{ij})$为实n阶方阵，其秩数等于n。试给出 $D^{(n)}$的一个一一变换φ，使

$$\varphi:(0,\cdots,0,\overset{(j)}{1},0,\cdots,0)\longmapsto(a_{1j},a_{2j},\cdots,a_{nj}),\ j=1,2,\cdots,n.$$

并写出φ的逆变换。

3．考虑n阶方阵的集合$M_n(D)$与n维向量的集合 $D^{(n)}$ 的幂集$P(D^{(n)})$。试给出从$M_n(D)$到$P(D^{(n)})$的一个映射，并说明你给出的这个映射有何意义。

4．考虑 $M_n(F)$。规则

$$\varphi_1:(a_{ij})\longmapsto(b_{ij}),\quad b_{ij}=\frac{a_{ij}+a_{ji}}{2},$$

$$\varphi_2:(a_{ij})\longmapsto(c_{ij}),\quad c_{ij}=\frac{a_{ij}-a_{ji}}{2},$$

试说明φ_1与φ_2都是$M_n(F)$ 的变换。它们是不是单变换？是不是满变换？为什么？并求$M_n(F)$的子集：

$$S_1=\{(a_{ij})\in M_n(F)\mid\varphi_1((a_{ij}))=(a_{ij})\},$$
$$S_2=\{(a_{ij})\in M_n(F)\mid\varphi_2((a_{ij}))=(a_{ij})\}.$$

5．设 $A=\{a,b\}$。给A建立两种代数运算。并指明这两个运算满足什么算律。

6．关于整数集Z的规则

$$\varphi:(a,b)\longmapsto a+b-ab$$

是Z的一个代数运算。φ是否满足结合律，交换律？

7. 考虑关于$D[x]$的下列规则

$\sim_1$: $f(x)\sim_1 g(x)$ 当且仅当$\dfrac{f(x)}{(f(x),f'(x))}=\dfrac{g(x)}{(g(x),g'(x))}$;

$\sim_2$: $f(x)\sim_2 g(x)$当且仅当$f(x)(g(x),\ g'(x))=g(x)(f(x),\ f'(x))$;

$\sim_3$: 如果$f(x)$, $g(x)$的首系数分别为a_0, b_0那么

$$f(x)\sim_3 g(x)\ \text{当且仅当}\frac{f(x)}{a_0(f(x),f'(x))}=\frac{g(x)}{b_0(g(x),\ g'(x))},$$

如果$f(x)=0$ 或 $g(x)=0$，那么

$$f(x)\sim_3 g(x)\ \text{当且仅当}\ f(x)=g(x)=0.$$

试述以上三个规则哪个是$D[x]$的关系，哪个是$D[x]$的等价关系？写出由等价关系所决定的分类。

8. 考虑n维向量集合$D^{(n)}$的幂集 $P(D^{(n)})$。

1）给出 $P(D^{(n)})$ 的一个分类Σ，并指出由Σ所决定的等价关系。

2）给出 $P(D^{(n)})$的一个等价关系$\sim$，并指出由$\sim$所决定的分类。

第八章　矩阵的运算

在线性方程组的讨论中已经充分地显示出矩阵的重要作用。矩阵在解线性方程组中所发挥的作用主要是通过矩阵的初等变换来实现的。此外，矩阵还能进行代数运算。正是由于矩阵有了运算方法，又使得它在许多方面有着重要的应用，成为非常有用的工具。本章就来讨论矩阵的代数运算。注意，所讨论的矩阵都是某个取定的数域F上的矩阵，所提到的数都是同一个数域F中的数。

§1　矩阵的运算及其性质

加法　设

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},\quad B=\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\cdots&\cdots&\cdots&\cdots\\b_{m1}&b_{m2}&\cdots&b_{mn}\end{pmatrix}$$

为任二$m\times n$矩阵。称 $m\times n$ 矩阵

$$C=\begin{pmatrix}a_{11}+b_{11}&a_{12}+b_{12}&\cdots&a_{1n}+b_{1n}\\a_{21}+b_{21}&a_{22}+b_{22}&\cdots&a_{2n}+b_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{m1}+b_{m1}&a_{m2}+b_{m2}&\cdots&a_{mn}+b_{mn}\end{pmatrix}$$

为A与B的和，记作 $C=A+B$。求和的方法叫做加法。

例如，若

$$A=\begin{pmatrix}-1&0&7\\2&5&-3\end{pmatrix},\quad B=\begin{pmatrix}6&4&4\\-8&0&1\end{pmatrix}$$

则
$$A+B=\begin{pmatrix}-1&0&7\\2&5&-3\end{pmatrix}+\begin{pmatrix}6&4&4\\-8&0&1\end{pmatrix}=\begin{pmatrix}5&4&11\\-6&5&-2\end{pmatrix}.$$

要注意的是：只有在两个矩阵的行数和列数都分别相等时，才

能相加。

数乘矩阵　　**对任$m\times n$矩阵**

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}$$

k为任一数，称$m\times n$矩阵

$$D=\begin{pmatrix}ka_{11} & ka_{12} & \cdots & ka_{1n}\\ ka_{21} & ka_{22} & \cdots & ka_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ ka_{m1} & ka_{m2} & \cdots & ka_{mn}\end{pmatrix}$$

为数k与矩阵A的纯量积，记作$D=kA$，或$D=Ak$。求纯量积的方法叫做倍数运算。

上面这两种运算都很简单，实际上矩阵的加法就是直接通过诸元素的加法来进行的，倍数运算就是直接通过诸元素的倍数来进行的。

关于矩阵加法和倍数运算具有下列性质：

1）**加法交换律**：设A、B为任二$m\times n$矩阵，则

$A+B=B+A$。

2）**加法结合律**：设A、B、C是任三个 $m\times n$矩阵，则

$(A+B)+C=A+(B+C)$。

3）**有零矩阵**：

$$0=\begin{pmatrix}0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 0\end{pmatrix},$$

这是一个元素全为零的$m\times n$矩阵，它具有性质：

$A+0=A=0+A$，

其中A为任一$m\times n$矩阵。

4）对任一$m\times n$矩阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{2n} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix},$$

存在$m\times n$矩阵

$$-A=\begin{pmatrix} -a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & -a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{m1} & -a_{m2} & \cdots & -a_{mn} \end{pmatrix}$$

使得

$$A+(-A)=(-A)+A=0.$$

我们把矩阵$-A$叫做A的负矩阵，把 $A+(-B)$叫做A与B的差，记作 $A-B=A+(-B)$。求两个矩阵之差的方法叫做减法。

5) $(k+l)A=kA+lA$;

6) $k(A+B)=kA+kB$;

7) $(kl)A=k(lA)$;

8) $1\cdot A=A$.

其中5)—8)里的A,B为任二$m\times n$矩阵。

除了以上两种运算外，矩阵还能进行乘法运算。正是看来有些玄妙的乘法运算才使矩阵成为在许多方面有着重要应用的数学工具。说是有些玄妙的矩阵乘法其实并不玄妙，它也是客观规律的一种反映。例如，我们考虑两个二元线性代数式：

$$\begin{cases} y_1=a_{11}x_1+a_{12}x_2 \\ y_2=a_{21}x_1+a_{22}x_2 \end{cases} \quad (1)$$

与

$$\begin{cases} z_1=b_{11}y_1+b_{12}y_2 \\ z_2=b_{21}y_1+b_{22}y_2. \end{cases} \quad (2)$$

在分析学中，这是两个二元线性（齐次）函数；在代数学中，它们就是两个变量的线性替换（或叫做变数的线性变换）；而在解析几何里（系数 a_{ij}，b_{ij}满足一定条件时）可以表示两个坐标旋转。我

们把（1）代入（2）即有

$$z_1 = b_{11}(a_{11}x_1 + a_{12}x_2) + b_{12}(a_{21}x_1 + a_{22}x_2)$$
$$z_2 = b_{21}(a_{11}x_1 + a_{12}x_2) + b_{22}(a_{21}x_1 + a_{22}x_2),$$

将右端整理后，得

$$\begin{cases} z_1 = (b_{11}a_{11} + b_{12}a_{21})x_1 + (b_{11}a_{12} + b_{12}a_{22})x_2 \\ z_2 = (b_{21}a_{11} + b_{22}a_{21})x_1 + (b_{21}a_{12} + b_{22}a_{22})x_2. \end{cases} \quad (3)$$

（3）式说明：如果把（1），（2）看成两个二元线性函数，那么它们的复合函数还是线性（齐次）的，并且其间的系数关系如下：

线性函数（1）的系数　　线性函数（2）的系数

$$a_{11}\ a_{12} \qquad\qquad b_{11}\ b_{12}$$
$$a_{21}\ a_{22} \qquad\qquad b_{21}\ b_{22}$$

复合函数（3）的系数

$$b_{11}a_{11} + b_{12}a_{21} \quad b_{11}a_{12} + b_{12}a_{22}$$
$$b_{21}a_{11} + b_{22}a_{21} \quad b_{21}a_{12} + b_{22}a_{22}.$$

如果把（1）、（2）看做是变数的线性变换，那么这两次变换的合成还是一个变数的线性变换，并且合成变换的系数与原变换的系数之间有如上的联系。再如（1）、（2）表示两个坐标旋转时，旋转角分别为 θ_1，θ_2，那么连续施行两次旋转的结果也是一个旋转且旋转角设为θ_3，这时它们系数间的上述联系表明：$\theta_3 = \theta_1 + \theta_2$。

上述事实表明，如果去掉（1）、（2）的各种具体解释，就其抽象的统一的规律而言恰好是两个二阶方阵合成得到一个确定的二阶方阵的一般方法：

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \to \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ a_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{pmatrix},$$

这就给我们建立矩阵的乘法提供了一个基础。

乘法　设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

分别为$m\times n$矩阵和$n\times p$矩阵，我们把$m\times p$矩阵

$$C=\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21}+\cdots+a_{1n}b_{n1} & a_{11}b_{12}+a_{12}b_{22}+\cdots+a_{1n}b_{n2} & \cdots & a_{11}b_{1p}+a_{12}b_{2p}+\cdots+a_{1n}b_{np} \\ a_{21}b_{11}+a_{22}b_{21}+\cdots+a_{2n}b_{n1} & a_{21}b_{12}+a_{22}b_{22}+\cdots+a_{2n}b_{n2} & \cdots & a_{21}b_{1p}+a_{22}b_{2p}+\cdots+a_{2n}b_{np} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}b_{11}+a_{m2}b_{21}+\cdots+a_{mn}b_{n1} & a_{m1}b_{12}+a_{m2}b_{22}+\cdots+a_{mn}b_{n2} & \cdots & a_{m1}b_{1p}+a_{m2}b_{2p}+\cdots+a_{mn}b_{np} \end{pmatrix}$$

叫做A与B的积，记作 $C=AB$. 求两个矩阵积的方法叫做矩阵乘法。

利用和号Σ可以把积矩阵C写得集中一些：

$$C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix}$$

$$c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}, \quad i=1,2,\cdots,m; \quad j=1,2,\cdots,p.$$

要注意，只有第一个矩阵A的列数与第二个矩阵B的行数相同时，A与B才能相乘，并且积矩阵C的行数与第一个矩阵A的行数相同，列数与第二个矩阵B的列数相同。例如，

$$\begin{pmatrix} 1 & 3 & 2 & 5 \\ 4 & -1 & 0 & 1 \\ 0 & 1 & -1 & 22 \end{pmatrix}\begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & -1 & 1 \\ 0 & 1 & 7 & 5 \\ 3 & 2 & 1 & 0 \end{pmatrix}=\begin{pmatrix} 19 & 19 & 18 & 14 \\ 6 & 4 & 10 & 3 \\ 67 & 45 & 14 & -4 \end{pmatrix},$$

$$(1\ \ 2\ \ 3\ \ 4\ \ 0)\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 0 \end{pmatrix}=(30), \quad \begin{pmatrix} 1 & 0 & 3 \\ 2 & -1 & 0 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -2 & 4 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 3 & 2 \\ 8 & -6 \end{pmatrix}.$$

矩阵乘法有以下重要性质：

9）**乘法结合律**：设A、B、C分别为$m\times n$、$n\times p$、$p\times q$ 矩阵，则 $(AB)C=A(BC)$。

10）**乘法对加法的分配律**：设 A、B为 $m\times n$矩阵，C为$n\times p$矩阵，D为 $q\times m$ 矩阵，则

$$(A+B)C=AC+BC,\quad D(A+B)=DA+DB.$$

11） $(kA)B=A(kB)=k(AB)$，

其中A、B分别为$m\times n$，$n\times p$ 矩阵，k 为任意数。

我们只给出结合律的证明，其它两条的证明是容易的，建议读者补上其证明。

现在来证明结合律．令

$$A=\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \cdots & \cdots & \cdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix},\quad B=\begin{pmatrix}b_{11} & \cdots & b_{1p}\\ \cdots & \cdots & \cdots\\ b_{n1} & \cdots & b_{np}\end{pmatrix},\quad C=\begin{pmatrix}c_{11} & \cdots & c_{1q}\\ \cdots & \cdots & \cdots\\ c_{p1} & \cdots & c_{pq}\end{pmatrix}.$$

于是

$$(AB)C=\left[\begin{pmatrix}a_{11} & \cdots & a_{1n}\\ \cdots & \cdots & \cdots\\ a_{m1} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}b_{11} & \cdots & b_{1p}\\ \cdots & \cdots & \cdots\\ b_{n1} & \cdots & b_{np}\end{pmatrix}\right]\begin{pmatrix}c_{11} & \cdots & c_{1q}\\ \cdots & \cdots & \cdots\\ c_{p1} & \cdots & c_{pq}\end{pmatrix}$$

$$=\begin{pmatrix}\sum_{l=1}^{n}a_{1l}b_{l1} & \cdots & \sum_{l=1}^{n}a_{1l}b_{lp}\\ \cdots & \cdots & \cdots\\ \sum_{l=1}^{n}a_{ml}b_{l1} & \cdots & \sum_{l=1}^{n}a_{ml}b_{lp}\end{pmatrix}\begin{pmatrix}c_{11} & \cdots & c_{1q}\\ \cdots & \cdots & \cdots\\ c_{p1} & \cdots & c_{pq}\end{pmatrix}$$

$$=\begin{bmatrix}\sum_{t=1}^{p}\left(\sum_{l=1}^{n}a_{1l}b_{lt}\right)c_{t1} & \cdots & \sum_{t=1}^{p}\left(\sum_{l=1}^{n}a_{1l}b_{lt}\right)c_{tq}\\ \cdots & \cdots & \cdots\\ \sum_{t=1}^{p}\left(\sum_{l=1}^{n}a_{ml}b_{lt}\right)c_{t1} & \cdots & \sum_{t=1}^{p}\left(\sum_{l=1}^{n}a_{ml}b_{lt}\right)c_{tq}.\end{bmatrix}.$$

这是一个$m\times q$矩阵，它的第i行第j列元素为

$$d_{ij}=\sum_{t=1}^{p}\left(\sum_{l=1}^{n}a_{il}b_{lt}\right)c_{tj}=\sum_{t=1}^{p}\sum_{l=1}^{n}a_{il}b_{lt}c_{tj}.$$

又

$$A(BC)=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \cdots&\cdots&\cdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix}\left[\begin{pmatrix}b_{11}&\cdots&b_{1p}\\ \cdots&\cdots&\cdots\\ b_{n1}&\cdots&b_{np}\end{pmatrix}\begin{pmatrix}c_{11}&\cdots&c_{1q}\\ \cdots&\cdots&\cdots\\ c_{p1}&\cdots&b_{pq}\end{pmatrix}\right]$$

$$=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\ \cdots&\cdots&\cdots\\ a_{m1}&\cdots&a_{mn}\end{pmatrix}\begin{bmatrix}\sum_{t=1}^{p}b_{1t}c_{t1}&\cdots&\sum_{t=1}^{p}b_{1t}c_{tq}\\ \cdots&\cdots&\cdots\\ \sum_{t=1}^{p}b_{nt}c_{t1}&\cdots&\sum_{t=1}^{p}b_{nt}c_{tq}\end{bmatrix}$$

$$=\begin{bmatrix}\sum_{l=1}^{n}a_{1l}\left(\sum_{t=1}^{p}b_{lt}c_{t1}\right)&\cdots&\sum_{l=1}^{n}a_{1l}\left(\sum_{t=1}^{p}b_{lt}c_{tq}\right)\\ \cdots&\cdots&\cdots\\ \sum_{l=1}^{n}a_{ml}\left(\sum_{t=1}^{p}b_{lt}c_{t1}\right)&\cdots&\sum_{l=1}^{n}a_{ml}\left(\sum_{t=1}^{p}b_{lt}c_{tq}\right)\end{bmatrix}.$$

这也是一个$m\times q$矩阵，它的第i行第j列元素为

$$f_{ij}=\sum_{l=1}^{n}a_{il}\left(\sum_{t=1}^{p}b_{lt}c_{tj}\right)=\sum_{l=1}^{n}\sum_{t=1}^{p}a_{il}b_{lt}c_{tj}$$

$$=\sum_{t=1}^{p}\sum_{l=1}^{n}a_{il}b_{lt}c_{tj}.$$

由矩阵相等的定义可知：

$(AB)C=A(BC)$。证完。

最后需要指出，矩阵的乘法是把数的乘法，加法按照一定的规律结合起来进行的。它远比矩阵的加法、数与矩阵纯量乘法要复杂得多。因此，我们所熟悉的关于数的乘法的一些运算规律不能直接认为对矩阵乘法都是适用的。除了已经证明过的上述三条之外，乘法交换律对矩阵来说就不成立。

事实上，若A为$m\times n$矩阵，B为$n\times p$矩阵，这时AB为$m\times p$矩阵，在$m\neq p$时，B和A不能相乘，更谈不上$BA=AB$；即使$m=p$时，则AB是m阶方阵，而BA是n阶方阵，在$m\neq n$时，也谈不上AB

是否等于BA。由此可见，只有在$m=n=p$时AB与BA才可能相同，但就是在这种情形下，乘法也不一定可交换。例如，取

$$A=\begin{pmatrix}0&0\\0&1\end{pmatrix},\quad B=\begin{pmatrix}0&0\\1&0\end{pmatrix}$$

则

$$AB=\begin{pmatrix}0&0\\0&1\end{pmatrix}\begin{pmatrix}0&0\\1&0\end{pmatrix}=\begin{pmatrix}0&0\\1&0\end{pmatrix},$$

$$BA=\begin{pmatrix}0&0\\1&0\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}.$$

按矩阵相等的定义，$AB\neq BA$。这里也顺便说明了：两个都不为零的矩阵乘积可能为零。

最重要的矩阵是方阵，我们用M_n表示数域F上一切n阶方阵的集合。于是对M_n的元素可以进行加法、倍数及乘法运算，并且满足上述1）—11）条全部性质。和我们已知的全体整数、一元多项式的全体相比较，它们有着同样的运算方法（主要是指加法、乘法）并且满足同样的运算规律（只有乘法交换律除外）。所以它在代数学里有着重要的地位。

前面在第五章说过转置矩阵的概念，和矩阵运算联系起来考虑，容易指出常见的四条性质：

1）$(A')'=A$；

2）$(A+B)'=A'+B'$；

3）$(kA)'=kA'$；

4）$(AB)'=B'A'$。

前三条证明是容易的，留给读者补证。我们来证4）：设

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix},\quad B=\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1p}\\b_{21}&b_{22}&\cdots&b_{2p}\\\cdots&\cdots&\cdots&\cdots\\b_{n1}&b_{n2}&\cdots&b_{np}\end{pmatrix}$$

$$AB=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}\begin{pmatrix}b_{11}&b_{12}&\cdots&b_{1p}\\b_{21}&b_{22}&\cdots&b_{2p}\\\cdots&\cdots&\cdots&\cdots\\b_{n1}&b_{n2}&\cdots&b_{np}\end{pmatrix}$$

$$
= \begin{pmatrix}
\sum_{l=1}^{n} a_{1l}b_{l1} & \sum_{l=1}^{n} a_{1l}b_{l2} \cdots & \sum_{l=1}^{n} a_{1l}b_{lp} \\
\sum_{l=1}^{n} a_{2l}b_{l1} & \sum_{l=1}^{n} a_{2l}b_{l2} \cdots & \sum_{l=1}^{n} a_{2l}b_{lp} \\
\cdots\cdots & \cdots\cdots & \cdots\cdots \\
\sum_{l=1}^{n} a_{ml}b_{l1} & \sum_{l=1}^{n} a_{ml}b_{l2} \cdots & \sum_{l=1}^{n} a_{ml}b_{lp}
\end{pmatrix},
$$

$$
(AB)' = \begin{pmatrix}
\sum_{l=1}^{n} a_{1l}b_{l1} & \sum_{l=1}^{n} a_{2l}b_{l1} & \cdots \sum_{l=1}^{n} a_{ml}b_{l1} \\
\sum_{l=1}^{n} a_{1l}b_{l2} & \sum_{l=1}^{n} a_{2l}b_{l2} & \cdots \sum_{l=1}^{n} a_{ml}b_{l2} \\
\cdots\cdots & \cdots\cdots & \cdots\cdots \\
\sum_{l=1}^{n} a_{1l}b_{lp} & \sum_{l=1}^{n} a_{2l}b_{lp} & \cdots \sum_{l=1}^{n} a_{ml}b_{lp}
\end{pmatrix}.
$$

而

$$
B'A' = \begin{pmatrix}
b_{11} & b_{21} & \cdots & b_{n1} \\
b_{12} & b_{22} & \cdots & b_{n2} \\
\cdots & \cdots & \cdots & \cdots \\
b_{1p} & b_{2p} & \cdots & b_{np}
\end{pmatrix}
\begin{pmatrix}
a_{11} & a_{21} & \cdots & a_{m1} \\
a_{12} & a_{22} & \cdots & a_{m2} \\
\cdots & \cdots & \cdots & \cdots \\
a_{1n} & a_{2n} & \cdots & a_{mn}
\end{pmatrix},
$$

$$
= \begin{pmatrix}
\sum_{l=1}^{n} a_{1l}b_{l1} & \sum_{l=1}^{n} a_{2l}b_{l1} & \cdots \sum_{l=1}^{n} a_{ml}b_{l1} \\
\sum_{l=1}^{n} a_{1l}b_{l2} & \sum_{l=1}^{n} a_{2l}b_{l2} & \cdots \sum_{l=1}^{n} a_{ml}b_{l2} \\
\cdots\cdots & \cdots\cdots & \cdots\cdots \\
\sum_{l=1}^{n} a_{1l}b_{lp} & \sum_{l=1}^{n} a_{2l}b_{lp} & \cdots \sum_{l=1}^{n} a_{ml}b_{lp}
\end{pmatrix}.
$$

于是根据矩阵相等的定义，必有 $(AB)' = B'A'$。证完。

最后我们来介绍分块矩阵的运算。在具体对矩阵运算时，往往

把矩阵分成一些小块。对矩阵作了适当的分块后，在进行运算时，就可把每个小块看作是一个元素再作同样的运算。

进行分块的方法是对一个矩阵A从上到下把行加以分组，从左至右把列加以分组，这样就把矩阵A划分成一些小矩阵。例如，设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

为$m\times n$矩阵，作分块如下：

$$A=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \cdots & \cdots & \cdots & \cdots \\ A_{p1} & A_{p2} & \cdots & A_{pq} \end{pmatrix}\begin{matrix} (\text{共}m_1\text{行}) \\ (\text{共}m_2\text{行}) \\ \vdots \\ (\text{共}m_p\text{行}) \end{matrix}$$
$$(\text{共}n_1\text{列})\ (\text{共}n_2\text{列})\cdots(\text{共}n_q\text{列})$$

其中$m_1+m_2+\cdots+m_p=m$，$n_1+n_2+\cdots+n_q=n$，A_{ij}是$m_i\times n_j$矩阵，叫做A的子块，通常把分为子块的阵叫做**分块阵**。例如

$$A=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix},\quad B=\begin{pmatrix} 1 & -1 & -1 & -1 & 0 \\ 2 & 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & -1 & -3 \\ 2 & 3 & 1 & -1 & 1 \end{pmatrix}.$$

将A、B分块为

$$A=\left(\begin{array}{cc:cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hdashline -1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right)=\begin{pmatrix} I_2 & O \\ A_1 & I_2 \end{pmatrix},$$

$$B=\left(\begin{array}{ccc:cc} 1 & -1 & -1 & -1 & 0 \\ \hdashline 2 & 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & -1 & -3 \\ \hdashline 2 & 3 & 1 & -1 & 1 \end{array}\right)=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix}$$

下面来介绍分块矩阵的运算。

加法 设 A，B 均为 $m\times n$矩阵，将 A, B 用同样的分法分成分块阵：

$$A=\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1q}\\ A_{21} & A_{22} & \cdots & A_{2q}\\ \cdots & \cdots & \cdots & \cdots\\ A_{p1} & A_{p2} & \cdots & A_{pq}\end{pmatrix},\quad B=\begin{pmatrix}B_{11} & B_{12} & \cdots & B_{1q}\\ B_{21} & B_{22} & \cdots & B_{2q}\\ \cdots & \cdots & \cdots & \cdots\\ B_{p1} & B_{p2} & \cdots & B_{pq}\end{pmatrix}$$

其中 A_{ij}、B_{ij} 均为 $m_i\times n_j$ 矩阵。那么

$$A+B=\begin{pmatrix}A_{11}+B_{11} & A_{12}+B_{12} & \cdots & A_{1q}+B_{1q}\\ A_{21}+B_{21} & A_{22}+B_{22} & \cdots & A_{2q}+B_{2q}\\ \cdots & \cdots & \cdots & \cdots\\ A_{p1}+B_{p1} & A_{p2}+B_{p2} & \cdots & A_{pq}+B_{pq}\end{pmatrix}.$$

数乘矩阵 设A为$m\times n$矩阵，将A分成分块矩阵

$$A=\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1q}\\ A_{21} & A_{22} & \cdots & A_{2q}\\ \cdots & \cdots & \cdots & \cdots\\ A_{p1} & A_{p2} & \cdots & A_{pq}\end{pmatrix}.$$

k 为任一数。那么

$$kA=\begin{pmatrix}kA_{11} & kA_{12} & \cdots & kA_{1q}\\ kA_{21} & kA_{22} & \cdots & kA_{2q}\\ \cdots & \cdots & \cdots & \cdots\\ kA_{p1} & kA_{p2} & \cdots & kA_{pq}\end{pmatrix},$$

其中每个 kA_{ij} 也是数k与矩阵 A_{ij} 的纯量积。

乘法 设 A，B分别为 $m\times n$矩阵，$n\times t$ 矩阵。对 A，B分块，这时要求A的列的分法与B的行的分法相一致。即

$$A=\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1q}\\ A_{21} & A_{22} & \cdots & A_{2q}\\ \cdots & \cdots & \cdots & \cdots\\ A_{p1} & A_{p2} & \cdots & A_{pq}\end{pmatrix},$$

（共n_1列）（共n_2列）$\cdots$（共n_q列）

$$B=\begin{pmatrix}B_{11} & B_{12} & \cdots & B_{1s}\\ B_{21} & B_{22} & \cdots & B_{2s}\\ \cdots & \cdots & \cdots & \cdots\\ B_{q1} & B_{q2} & \cdots & B_{qs}\end{pmatrix}\begin{matrix}(\text{共}n_1\text{行})\\(\text{共}n_2\text{行})\\ \vdots\\(\text{共}n_q\text{行})\end{matrix}.$$

那么

$$AB=\begin{pmatrix}C_{11} & C_{12} & \cdots & C_{1s}\\ C_{21} & C_{22} & \cdots & C_{2s}\\ \cdots & \cdots & \cdots & \cdots\\ C_{p1} & C_{p2} & \cdots & C_{ps}\end{pmatrix},$$

其中 $c_{ij}=A_{i1}B_{1j}+A_{i2}B_{2j}+\cdots+A_{iq}B_{qj}$ $i=1,2,\cdots,p$, $j=1,2,\cdots,s$。

以上用分块矩阵作的加法、纯量乘法和乘法运算的结果都可以用矩阵的加法、纯量乘法和乘法的定义来验证它们的正确性，这里不详细说明了。

将矩阵A分块后，如果它的形式如

$$A=\begin{pmatrix}A_{11} & A_{12} & \cdots & A_{1p}\\ 0 & A_{22} & \cdots & A_{2p}\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & A_{pp}\end{pmatrix} \text{或} A=\begin{pmatrix}A_{11} & 0 & \cdots & 0\\ A_{21} & A_{22} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ A_{p1} & A_{p2} & \cdots & A_{pp}\end{pmatrix},$$

则称前者为准上三角形矩阵，后者为准下三角形矩阵。在A_{11}, A_{22}, …, A_{pp}都是方阵时，显然A也是方阵。这时由 Laplace 展开式可知

$$\det A=\det A_{11}\det A_{22}\cdots\det A_{pp}.$$

特别地，若A分块成以下形式

$$A=\begin{pmatrix}A_{11} & 0 & \cdots & 0\\ 0 & A_{22} & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & A_{pp}\end{pmatrix}$$

则称A为准对角形矩阵。这时有

$$\operatorname{rank}A=\operatorname{rank}A_{11}+\operatorname{rank}A_{22}+\cdots+\operatorname{rank}A_{pp}.$$

下面举例说明“准”的概念与分法有关。

例如，将矩阵

$$\begin{pmatrix}1 & 2 & 0 & 0\\ 3 & 8 & 0 & 0\\ 0 & 0 & 1 & 3\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 1\end{pmatrix}$$

按下列方法分块：

$$\left(\begin{array}{cc|cc}1 & 2 & 0 & 0\\ 3 & 8 & 0 & 0\\ \hline 0 & 0 & 1 & 3\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 1\end{array}\right),\quad \left(\begin{array}{cc|c|c}1 & 2 & 0 & 0\\ 3 & 8 & 0 & 0\\ 0 & 0 & 1 & 3\\ \hline 0 & 0 & 2 & 1\\ \hline 0 & 0 & 0 & 1\end{array}\right),\quad \left(\begin{array}{cc|c|c}1 & 2 & 0 & 0\\ \hline 3 & 8 & 0 & 0\\ \hline 0 & 0 & 1 & 3\\ 0 & 0 & 2 & 1\\ 0 & 0 & 0 & 1\end{array}\right)$$

则它们分别为准对角形、准上三角形、准下三角形矩阵。由此可见，在具体运用分块矩阵运算时，要充分考虑问题选取何种分法最合适。

我们以后将会看到，把矩阵分块进行运算有许多方便之处。比如，利用矩阵分块运算，可以把数的等式，向量的等式和矩阵的等式统一起来：

线性方程组的一般写法为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{cases} \quad (4)$$

如果令

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

则有与（4）等效的写法为

$$AX = B. \quad (5)$$

再如令

$$\alpha_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \alpha_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \cdots, \quad \alpha_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}, \quad \beta = B,$$

则又有与（4）、（5）等效写法为

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = \beta. \quad (6)$$

这样数的等式（4）、矩阵的等式（5）、向量的等式（6）就成为一个问题的三种等效的表现形式。

末了引进倍数与幂的记号：把 $\underbrace{A + A + \cdots + A}_{m个}$ 记作 mA，把 $\underbrace{A \cdot A \cdot \cdots \cdot A}_{m个}$ 记作 A^m，这里 m 为正整数。

练　习　一

1．计算

1）$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 1 & -2 \\ 3 & -2 & 1 \\ -3 & 1 & -1 \end{pmatrix}$，

2）$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 1 & 4 \\ 2 & 0 & 3 & 1 \end{pmatrix} + \begin{pmatrix} 3 & 2 & 1 & 0 \\ 2 & -1 & -1 & 1 \\ 0 & -1 & 3 & 2 \end{pmatrix} - \begin{pmatrix} -1 & 2 & 3 & 4 \\ 0 & 2 & 0 & -1 \\ -1 & 1 & 3 & 1 \end{pmatrix}$。

2．求X，使

$$\begin{pmatrix} 2 & 1 & 1 \\ 3 & 1 & 2 \\ -1 & 0 & 1 \end{pmatrix} + X - \begin{pmatrix} 2 & 3 & 0 \\ -1 & 0 & -1 \\ 2 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -3 & -1 & 2 \end{pmatrix}.$$

3．计算

1）$\begin{pmatrix} 1 & 3 & -2 & 1 \\ 2 & 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \\ 3 \\ 1 \end{pmatrix}$，　　2）$(a_1 a_2 \cdots a_n)\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$，

3）$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}(b_1 b_2 \cdots b_n)$，4）$(x \ \ y \ \ 1)\begin{pmatrix} a & b & d \\ b & c & e \\ d & e & f \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$。

4．设

1）$A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}$，$B = \begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$。

2）$A = \begin{pmatrix} a & b & c \\ c & b & a \\ 1 & 1 & 1 \end{pmatrix}$，$B = \begin{pmatrix} 1 & a & c \\ 1 & b & b \\ 1 & c & a \end{pmatrix}$，

求$AB - BA$。

5．计算

1）$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$，　2）$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}^n$，

3) $\begin{pmatrix}1 & 1 & 0\\0 & 1 & 1\\0 & 0 & 1\end{pmatrix}^n$.

6. 设

$$A=\begin{pmatrix}1 & 2\\2 & 1\\3 & 1\end{pmatrix},$$

求AA'，$A'A$。

7. 设 $f(\lambda)=a_0\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m$，$A$是一个$n\times n$矩阵。定义

$$f(A)=a_0A^m+a_1A^{m-1}+\cdots+a_{m-1}A+a_mI_n.$$

设

1) $f(\lambda)=\lambda^2-\lambda+1$，　$A=\begin{pmatrix}2 & 1 & 1\\3 & 1 & 2\\1 & -1 & 0\end{pmatrix}$；

2) $f(\lambda)=\lambda^2-5\lambda+3$，$A=\begin{pmatrix}2 & -1\\-3 & 3\end{pmatrix}$，

试求 $f(A)$。

8. 设

$$A=\begin{pmatrix}\alpha & 1 & 0 & 0\\0 & \alpha & 0 & 0\\0 & 0 & \beta & 1\\0 & 0 & 1 & \beta\end{pmatrix},\quad B=\begin{pmatrix}\alpha & 0 & 0 & 0\\1 & \alpha & 0 & 0\\0 & 0 & \beta & 0\\0 & 0 & 1 & \beta\end{pmatrix},$$

求 A^2，ABA，B^3。

9. 设 A,B都是 $m\times n$矩阵。证明：方程

$$A+X=B$$

有且只有一解。

10. 证明：$kA=0$ 必要而且只要$k=0$ 或$A=0$。

11. 设

1) $A=\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}$，2) $A=\begin{pmatrix}0 & 1 & 0\\0 & 0 & 1\\0 & 0 & 0\end{pmatrix}$。

求所有与A可交换的矩阵。

12．试问：对任二n阶方阵A与B，等式

$$det(A+B)=detA+detB$$

总成立吗？

§2 可逆矩阵

上一节对加法运算，我们定义了加法的逆运算——减法。类似地，自然要考虑矩阵乘法的逆运算问题。为此首先指出，对于n阶方阵的乘法来说，也存在一个和数1有同样作用的n阶方阵，即

$$I_n=\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 1\end{pmatrix}$$

使得

$$AI_n=I_nA=A,\quad A\in M_n,$$

我们称I_n为n阶单位阵，简称为单位阵。

进而我们给出

定义1　对于n阶方阵A，若存在n阶方阵B，使得

$$AB=BA=I_n \tag{1}$$

则称A为可逆矩阵。

容易指出，对n阶方阵A最多只能有一个n阶方阵B，使（1）式成立。

事实上，若还有n阶方阵C，有

$$AC=CA=I_n. \tag{2}$$

我们说，必有$B=C$。因为，一方面

$$BAC=(BA)C=I_nC=C,$$

另一方面

$$BAC=B(AC)=BI_n=B,$$

所以

$$B=C.$$

我们称满足（1）式的那个唯一的 B 叫做 A 的逆矩阵，记作 A^{-1}，亦即

$$AA^{-1}=A^{-1}A=I_n.$$

必须强调指出的是，并不是每个方阵都有逆矩阵，即并不是每个方阵都是可逆矩阵。

例如，我们看二阶方阵

$$A=\begin{pmatrix}0&0\\0&1\end{pmatrix},\quad D=\begin{pmatrix}0&0\\1&0\end{pmatrix}$$

有

$$DA=\begin{pmatrix}0&0\\1&0\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}. \qquad (3)$$

于是，我们可以断言 A 没有逆矩阵。如其不然，则有 A^{-1}，使

$$AA^{-1}=I_2.$$

那么将（3）式两端从右边乘以 A^{-1}，得

$$(DA)A^{-1}=OA^{-1},\ \text{即}D=0.$$

这是不可能的，从而 A 不能有逆矩阵，即 A 不可逆。因此，矩阵乘法没有逆运算。

既然不是每个方阵都可逆，这自然有以下问题：什么样的方阵是可逆的？如果 A 可逆，如何求 A 的逆矩阵？为了解决这些问题，我们给出

定义2　设 A_{ij} 为矩阵

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$$

中元素 a_{ij} 的代数余子式。矩阵

$$A^*=\begin{pmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\cdots&\cdots&\cdots&\cdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{pmatrix}$$

称为 A 的伴随矩阵。

根据行列式的展开定理，可得到

$$AA^* = A^*A = \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & |A| \end{pmatrix} = |A|\, I_n. \tag{4}$$

例如　设

$$A = \begin{pmatrix} 1 & -4 & -3 \\ 2 & -5 & -3 \\ -1 & 2 & 1 \end{pmatrix},$$

求A^*.

解　因为

$$A_{11} = 1,\quad A_{12} = 1,\quad A_{13} = -1,$$
$$A_{21} = -2,\quad A_{22} = -2,\quad A_{23} = 2,$$
$$A_{31} = -3,\quad A_{32} = -3,\quad A_{33} = 3.$$

故

$$A^* = \begin{pmatrix} 1 & -2 & -3 \\ 1 & -2 & -3 \\ -1 & 2 & 3 \end{pmatrix}.$$

请读者验证一下

$$AA^* = A^*A = |A|\, I_3 = 0\ .$$

定理　n阶方阵 $A = (a_{ij})$是可逆的充分必要条件是：A的行列式 $|A| \neq 0$，而且当A可逆时，

$$A^{-1} = \frac{1}{|A|}A^*. \tag{5}$$

证明　先证充分性．即设 $|A| \neq 0$，去证A为可逆的．由（5）式可知，当 $|A| \neq 0$时，$\frac{1}{|A|}A^*$ 存在并且由（4）式知

$$A\left(\frac{1}{|A|}A^*\right) = \frac{1}{|A|}(AA^*) = \frac{1}{|A|}\,|A|\, I_n = I_n,$$

$$\left(\frac{1}{|A|}A^*\right)A = \frac{1}{|A|}(A^*A) = \frac{1}{|A|}\,|A|\, I_n = I_n$$

这说明存在$\frac{1}{|A|}A^*$，使（1）成立，故A为可逆的。

再证必要性。即若A为可逆的，去证$|A|\neq 0$。因为A是可逆的，有A^{-1}存在，令$A^{-1}=(\alpha_1,\alpha_2\cdots\alpha_n)$，$\alpha_i$为$n\times 1$矩阵，$i=1,2,\cdots,n$。于是由

$$AA^{-1}=I_n,\quad 即A(\alpha_1\alpha_2\cdots\alpha_n)=I_n$$

可得

$$A\alpha_1=\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}.$$

因为阵A的逆阵是唯一的，可知线性方程组

$$AX=\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}$$

有唯一解，从而其系数阵A的行列式$|A|\neq 0$。证完

由充分性的证明过程可知，当$|A|\neq 0$时，A的逆阵为$A^{-1}=\frac{1}{|A|}A^*$。

定理中的（5）式就是求逆矩阵的公式。

推论 **A为可逆的充分必要条件是A为满秩的。**

例如 判别下列矩阵是否可逆？如是，求其逆阵。

$$A=\begin{pmatrix}1&0&0\\2&1&0\\-3&2&1\end{pmatrix},\quad B=\begin{pmatrix}1&2&0&0\\-1&2&1&3\\0&0&2&4\\3&0&1&2\end{pmatrix}.$$

解 因为$|A|=1\neq 0$，所以由定理知，A可逆，并且由于

$A_{11}=1$，$A_{12}=-2$，$A_{13}=7$，

$A_{21}=0$，$A_{22}=1$，$A_{23}=-2$，

$A_{31}=0$，$A_{32}=0$，$A_{33}=1$。

故

$$A^{-1}=\begin{pmatrix}1&0&0\\-2&1&0\\7&-2&1\end{pmatrix}.$$

因为

$$|B|=\begin{vmatrix}1&2&0&0\\-1&2&1&3\\0&0&2&4\\3&0&1&2\end{vmatrix}=0,$$

由定理知，B不可逆。

下面来讨论可逆矩阵的性质。

命题 1 $(A^{-1})^{-1}=A$。

事实上，由于 $A^{-1}A=AA^{-1}=I_n$，不仅说明A^{-1}是A的逆阵，而且也说明A是A^{-1}的逆阵，也就是

$(A^{-1})^{-1}=A$。证完。

命题 2 若A是可逆阵，则转置阵A'也是可逆的，而且

$(A')^{-1}=(A^{-1})'$.

事实上，$A'(A^{-1})'=(A^{-1}A)'=I'_n=I_n$,

$(A^{-1})'A'=(AA^{-1})'=I_n'=I_n$.

故 $(A')^{-1}=(A^{-1})'$，证完。

命题 3 若A、B都可逆，则AB也可逆，而且

$(AB)^{-1}=B^{-1}A^{-1}$.

事实上，因为

$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AA^{-1}=I_n$,

$(B^{-1}A^{-1})(AB)=B^{-1}(A^{-1}A)B=B^{-1}B=I_n$.

故$(AB)^{-1}=B^{-1}A^{-1}$，证完。

命题 3 不难推广到s个可逆阵的情形：

设 A_1，A_2，…，A_s都是n阶可逆阵，那么 $A_1A_2\cdots A_s$ 也是可逆阵，而且

$(A_1A_2\cdots A_s)^{-1}=A_s^{-1}\cdots A_2^{-1}A_1^{-1}$.

命题 3 说明M_n中一切可逆阵的集合GL_n关于矩阵的乘法是封闭的。

最后，讨论一下矩阵方程：

$$AX=B.$$

如果A是一个可逆阵，那么这个方程有且只有一个解：

$$X=A^{-1}B. \qquad (6)$$

这里的B可以是$n\times m$矩阵，即：如果A是一个n阶可逆阵，B是一个$n\times m$矩阵，那么方程

$$AX=B$$

有且只有一个解

$$X=A^{-1}B,$$

这个解X也是一个$n\times m$矩阵。

特别地，当B是一个$n\times 1$矩阵时，公式（6）给出了克莱姆法则的另一个证明：

线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n.\end{cases} \qquad (7)$$

根据矩阵乘法定义，可以写成矩阵方程

$$AX=B, \qquad (8)$$

其中

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots a_{1n}\\ a_{21} & a_{22} & \cdots a_{2n}\\ \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots a_{nn}\end{pmatrix},\quad X=\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix},\quad B=\begin{pmatrix}b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}.$$

如果$|A|\neq 0$，那么A可逆，方程（8）有且只有一个解

$$X=A^{-1}B,$$

这也是方程组（7）的解。将$A^{-1}=\frac{1}{|A|}A^*$代入，乘出来就是克莱姆法则给出的公式。

如果A不是可逆阵，我们来讨论更一般的情形。设A是一个$p\times n$矩阵，B是一个$p\times m$矩阵，考虑方程

$$AX = B. \tag{9}$$

其中未知矩阵X是一个$n\times m$矩阵。现在用分块矩阵讨论这个问题：将B、X用分块阵写成

$$B=(B_1\ B_2\cdots B_m),\quad X=(X_1\ X_2\cdots X_m)$$

其中B_i都是$p\times 1$矩阵；X_i都是$n\times 1$矩阵。于是方程（9）可以写成

$$A(X_1X_2\cdots X_m)=(B_1B_2\cdots B_m).$$

它等价于方程组

$$\begin{cases}AX_1=B_1\\ AX_2=B_2\\ \cdots\cdots\cdots\cdots\\ AX_m=B_m.\end{cases} \tag{10}$$

再将A写成分块矩阵：

$$A=(A_1A_2\cdots A_n),$$

其中A_j是A的第j列，是$p\times 1$矩阵。于是方程组（10）可以写成

$$(A_1A_2\cdots A_n)X_k=B_k\quad(k=1,2,\cdots,m). \tag{11}$$

因为方程组（11）有解的条件是$B_k(k=1,2,\cdots,m)$可以被$A_1,A_2,\cdots,A_n$线性表出，所以方程（9）有解的充分必要条件是B的列向量组可以由A的列向量组线性表出。这个结论当然也包含A是可逆阵的情形。

方程$YA=B$也可以类似地加以讨论，这里就不详细讲了。

例如，求X，使

$$\begin{pmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{pmatrix}X=\begin{pmatrix}1&-1&2\\-1&3&4\\-2&0&5\end{pmatrix}.$$

解

$$X=\begin{pmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{pmatrix}^{-1}\begin{pmatrix}1&-1&2\\-1&3&4\\-2&0&5\end{pmatrix}$$

$$=\begin{pmatrix}-5&0&-8\\-3&-1&-6\\2&0&3\end{pmatrix}\begin{pmatrix}1&-1&2\\-1&3&4\\-2&0&5\end{pmatrix}$$

$$=\begin{pmatrix} 11 & 5 & -10 \\ 10 & 0 & -10 \\ -4 & -2 & 19 \end{pmatrix}.$$

练 习 二

1．判别下列矩阵是否可逆？如可逆，求出其逆阵．

1） $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，$ad-bc=1$， 2） $A=\begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 3 & 2 & -1 \end{pmatrix}$，

3） $A=\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \\ 3 & 4 & 3 \end{pmatrix}$， 4） $A=\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 7 & 2 & 3 \\ 2 & 5 & 1 & 2 \end{pmatrix}$

5） $A=\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$， $a_i\neq 0$，$i=1,2,\cdots,n$．

2．求满足下列条件的X：

1） $\begin{pmatrix} 1 & -5 \\ -1 & 4 \end{pmatrix}X=\begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$；

2） $X\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 2 & -3 \\ 2 & 0 & 4 \\ 0 & -1 & 5 \end{pmatrix}$；

3） $\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}X\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix}=\begin{pmatrix} 4 & 2 & 3 \\ 0 & -1 & 5 \\ 2 & 1 & 1 \end{pmatrix}$，

4） $\begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix}X=\begin{pmatrix} 2 & 5 \\ 1 & 4 \end{pmatrix}$．

3．

1）设n阶阵A,B都可逆，证明：$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$及$\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$都可逆，并且求出$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^{-1}$，$\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}^{-1}$；

2）计算

$$\begin{pmatrix}3 & -2 & 0 & 0\\5 & -3 & 0 & 0\\0 & 0 & 3 & 4\\0 & 0 & 1 & 1\end{pmatrix}^{-1}, \quad \begin{pmatrix}0 & 0 & 0 & 1 & 2\\0 & 0 & 0 & 2 & 3\\1 & 1 & 0 & 0 & 0\\0 & 1 & 1 & 0 & 0\\0 & 0 & 1 & 0 & 0\end{pmatrix}^{-1}.$$

4．设

$$A=\begin{pmatrix}-1 & 0 & 0\\1 & -1 & 0\\1 & 1 & -1\end{pmatrix},$$

试计算

$$(A+2I_3)^{-1}(A^2-4I_3)\text{及}(A+2I_3)^{-1}(A-2I_3).$$

5．若$A^k=0$，则 $(I_n-A)^{-1}=I_n+A+A^2+\cdots+A^{k-1}$.

6．试问：有没有不是方阵的矩阵A,B满足$AB=I$呢？

7．若A可逆，则$-A$也可逆，并求$-A$的逆阵.

8．若 $AB=BA$，且A为可逆阵，则$A^{-1}B=BA^{-1}$.

9．若A可逆，则A^*可逆，并求A^*的逆阵.

10．设n阶方阵A适合$A^2=A$，且$A\neq I_n$，则A不可逆.

§3 初等矩阵

本节讨论一类特殊的可逆矩阵，即初等矩阵．它在矩阵的理论中有独特的作用.

下面三种n阶方阵统称为初等矩阵：

1）**倍法矩阵**

$$D_i(k)=\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & 1 & & & \\ & & & k & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1\end{pmatrix}(i\text{行})\quad k\neq 0.$$

2）**消法矩阵**

$$P_{ij}(a)=\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1\cdots a&&\\&&\ddots\ \vdots&&\\&&1&&\\&&&\ddots&\\&&&&1\end{pmatrix}\begin{matrix}\\\\(i\text{行})\\\\(j\text{行}),\\\\\\\end{matrix}\quad(i<j);$$

或

$$P_{ji}(a)=\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&\vdots\ \ddots&&\\&&a\cdots 1&&\\&&&\ddots&\\&&&&1\end{pmatrix}\begin{matrix}\\\\(i\text{行})\\\\(j\text{行}),\\\\\\\end{matrix}\quad(i<j).$$

3）换法矩阵

$$C_{ij}=\begin{pmatrix}1&&&&&&&&\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&0&\cdots\cdots&1&&&\\&&&\vdots&1&\vdots&&&\\&&&\vdots&\ddots&\vdots&&&\\&&&\vdots&1&\vdots&&&\\&&&1&\cdots\cdots&0&&&\\&&&&&&1&&\\&&&&&&&\ddots&\\&&&&&&&&1\end{pmatrix}\begin{matrix}\\\\\\(i\text{行})\\\\\\\\(j\text{行}),\\\\\\\\\end{matrix}\quad(i<j).$$

显然，初等矩阵都是可逆的，且

$$D_i(k)^{-1}=D_i\left(\frac{1}{k}\right),\ P_{ij}(a)^{-1}=P_{ij}(-a),\ C_{ij}^{-1}=C_{ij}.$$

这说明倍法矩阵的逆阵还是倍法矩阵，消法矩阵的逆阵还是消法矩阵，换法矩阵的逆阵还是换法矩阵。

命题　**用初等矩阵左乘矩阵A，恰等于用同类的初等变换变A的行；用初等矩阵右乘矩阵A，恰等于用同类的初等变换变A的列。**

证明　设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{im} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots & a_{jm} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

为$n\times m$矩阵。于是

$$D_i(k)A=\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & & \\ & & & k & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{im} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots & a_{jm} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

$$=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{i1} & ka_{i2} & \cdots & ka_{im} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots & a_{jm} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix};$$

$$P_{ij}(a)A=\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \cdots & a & \\ & & & \ddots & \vdots & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{im} \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots & a_{jm} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

$$=\begin{pmatrix} a_{11} & a_{12} & \cdots a_{1m} \\ \cdots & \cdots & \cdots \\ a_{i1}+aa_{j1} & a_{i2}+aa_{j2} & \cdots a_{im}+aa_{jm} \\ \cdots & \cdots & \cdots \\ a_{j1} & a_{j2} & \cdots a_{jm} \\ \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots a_{nm} \end{pmatrix};$$

$$C_{ij}A=\begin{pmatrix}1&&&&&&&&\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&0&\cdots&1&&&\\&&&\vdots&1&&\vdots&&\\&&&\vdots&&\ddots&\vdots&&\\&&&\vdots&&1&\vdots&&\\&&&1&\cdots&0&&&\\&&&&&&1&&\\&&&&&&&\ddots&\\&&&&&&&&1\end{pmatrix}\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\\cdots&\cdots&\cdots&\cdots\\a_{i1}&a_{i2}&\cdots&a_{im}\\\cdots&\cdots&\cdots&\cdots\\a_{j1}&a_{j2}&\cdots&a_{jm}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nm}\end{pmatrix}$$

$$=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1m}\\\cdots&\cdots&\cdots&\cdots\\a_{j1}&a_{j2}&\cdots&a_{jm}\\\cdots&\cdots&\cdots&\cdots\\a_{i1}&a_{i2}&\cdots&a_{im}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nm}\end{pmatrix}.$$

类似地可验证右乘的情形。证完。

定理1 **n阶方阵A可逆的充分必要条件是：A 可表成若干个初等矩阵之积。**

证明 充分性。设

$$A=P_1\ P_2\cdots P_s,$$

其中 $P_i(i=1,2,\cdots,s)$ 都是初等矩阵。由于初等矩阵都是可逆的，所以其积也是可逆的，故A为可逆阵。

必要性。设A为可逆阵，则A为满秩的。于是A在相抵之下的标准形必为

$$I_n=\begin{pmatrix}1&&&\\&1&&\\&&\ddots&\\&&&1\end{pmatrix}.$$

这就是说，用一些适当的初等变换，可将A化为单位阵I_n。根据本节命题可知，每个初等变换的作用都可以代之以用同类的初等矩阵去左乘或右乘。于是，A相抵于单位阵I_n，可写成

$$P_s\cdots P_2P_1AQ_1Q_2\cdots Q_t=I_n,$$

其中 P_i, Q_j 都是初等矩阵，由于初等矩阵都是可逆矩阵，且其逆阵也是初等矩阵。因此上式又可写成

$$A = P_1^{-1}P_2^{-1}\cdots P_s^{-1}I_nQ_t^{-1}\cdots Q_2^{-1}Q_1^{-1}$$
$$= P_1^{-1}P_2^{-1}\cdots P_s^{-1}Q^{-1}_t\cdots Q_2^{-1}Q_1^{-1}.$$

即A表成了初等矩阵的乘积．证完．

推论　可逆矩阵可以只用若干次行的初等变换化为单位阵，也可以只用若干次列的初等变换化为单位阵．

证明　据上定理可知，可逆矩阵A可以表成初等矩阵之积，即

$$A = P_1P_2\cdots P_s.$$

将上式改写成

$$P_s^{-1}\cdots P_2^{-1}P_1^{-1}A = I_n.$$

这说明从A的左边依次乘一些适当初等矩阵，就将A化为单位阵了。而据本节命题知，左乘初等矩阵，相当于施行同类的行的初等变换，故A可经若干次行的初等变换化成单位阵．

类似地可证列的情形．证完．

这个推论给我们提供一个求逆矩阵的另一种方法．设A为n阶可逆阵．由推论，有

$$P_s\cdots P_2P_1A = I_n,\ \text{或}\ AQ_1Q_2\cdots Q_t = I_n.$$

于是

$$A^{-1} = P_s\cdots P_2P_1I_n,\ \text{或}\ A^{-1} = I_nQ_1Q_2\cdots Q_t.$$

其中 P_i，Q_j 为初等矩阵．这说明：如果用一系列的行（列）初等变换把可逆阵A化为单位阵，那么用同样的初等变换作用于单位阵I_n，就可得到A的逆矩阵 A^{-1}．从而就得到用初等变换求逆矩阵的方法：

作一个$n\times 2n$矩阵

$$(A\ \ I_n)$$

对这个矩阵施行行的初等变换，将它的左半部 A 化为单位阵．那么，右半部的单位阵I_n就同时化成了A^{-1}：

$$(A\ \ I_n) \longrightarrow (I_n\ \ A^{-1}).$$

例1 求

$$A=\begin{pmatrix}1&2&3\\2&1&2\\1&3&4\end{pmatrix}$$

的逆阵。

解 对$(A\ I_3)$施行行的初等变换：

$$\begin{pmatrix}1&2&3&1&0&0\\2&1&2&0&1&0\\1&3&4&0&0&1\end{pmatrix}$$

$$\xrightarrow{P_{21}(-2),P_{31}(-1)}\begin{pmatrix}1&2&3&1&0&0\\0&-3&-4&-2&1&0\\0&1&1&-1&0&1\end{pmatrix}$$

$$\xrightarrow{P_{23}(4)}\begin{pmatrix}1&2&3&1&0&0\\0&1&0&-6&1&4\\0&1&1&-1&0&1\end{pmatrix}$$

$$\xrightarrow{P_{12}(-2),P_{32}(-1)}\begin{pmatrix}1&0&3&13&-2&-8\\0&1&0&-6&1&4\\0&0&1&5&-1&-3\end{pmatrix}$$

$$\xrightarrow{P_{13}(-3)}\begin{pmatrix}1&0&0&-2&1&1\\0&1&0&-6&1&4\\0&0&1&5&-1&-3\end{pmatrix}.$$

于是

$$A^{-1}=\begin{pmatrix}-2&1&1\\-6&1&4\\5&-1&-3\end{pmatrix}.$$

例2 求

$$A=\begin{pmatrix}1&0&0\\2&1&0\\-3&2&1\end{pmatrix}$$

的逆阵.

解 作$(A\ I_3)$，并施行行的初等变换：

$$\begin{pmatrix}1&0&0&1&0&0\\2&1&0&0&1&0\\-3&2&1&0&0&1\end{pmatrix}$$

$$\xrightarrow{P_{21}(-2),P_{31}(3)}\begin{pmatrix}1&0&0&1&0&0\\0&1&0&-2&1&0\\0&2&1&3&0&1\end{pmatrix}$$

$$\xrightarrow{P_{32}(-2)}\begin{pmatrix}1&0&0&1&0&0\\0&1&0&-2&1&0\\0&0&1&7&-2&1\end{pmatrix}.$$

故

$$A^{-1}=\begin{pmatrix}1&0&0\\-2&1&0\\7&-2&1\end{pmatrix}.$$

定理 2 n阶方阵A可逆的充分必要条件是：

$$A=P_1\cdots P_s\begin{pmatrix}d_1&&&\\&d_2&&\\&&\ddots&\\&&&d_n\end{pmatrix},$$

其中P_i都是消法矩阵，$d_j\neq 0$，$i=1,2,\cdots,s$；$j=1,2,\cdots,n$。

证明 充分性。因为A等于诸可逆矩阵之积，所以A是可逆的。

必要性。设

$$A=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\cdots&\cdots&\cdots&\cdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}$$

为n阶可逆阵，据第四章§3命题可知，只用行的消法变换就可以把A化成三角形阵：

$$A\rightrightarrows\begin{pmatrix}d_1&a_{12}'&\cdots&a_{1n}'\\0&d_2&\cdots&a_{2n}'\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&d_n\end{pmatrix}.$$

由于A可逆，可知$d_1\neq 0$，$d_2\neq 0,\cdots,d_n\neq 0$。于是再用一些适当的行的消法变换就能把这个三角形阵化成对角形阵：

$$A\rightrightarrows\begin{pmatrix}d_1&a_{12}'&\cdots&a_{1n}'\\0&d_2&\cdots&a_{2n}'\\\cdots&\cdots&\cdots&\cdots\\0&0&\cdots&d_n\end{pmatrix}\rightrightarrows\begin{pmatrix}d_1&&&\\&d_2&&\\&&\ddots&\\&&&d_n\end{pmatrix}.$$

把每次施行的行的消法变换代之以相应的消法矩阵，便有

$$Q_s\cdots Q_1 A=\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}.$$

把诸消法矩阵Q_i移到右端，即得

$$A=Q_1^{-1}\cdots Q_s^{-1}\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

或

$$A=P_1\cdots P_s\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}.$$ 证完。

推论 若A是可逆矩阵，则

$$A=P_1\cdots P_t\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & d \end{pmatrix},$$

其中 $d=\det A$. 而$P_1,\cdots,P_t$为消法矩阵.

证明 据定理 2，首先指出用行的消法变换，将对角形可逆阵

$$\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

化成

$$\begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & d \end{pmatrix}.$$

因为

$$\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} \xrightarrow{P_{21}\left(\frac{1}{d_1}\right)} \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 1 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} \xrightarrow{P_{12}(1-d_1)}$$

$$\begin{pmatrix} 1 & (1-d_1)d_2 & 0 & \cdots & 0 \\ 1 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix} \xrightarrow{P_{21}(-1)} \begin{pmatrix} 1 & (1-d_1)d_2 & 0 & \cdots & 0 \\ 0 & d_1d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}$$

$$\xrightarrow{P_{12}\left(\frac{-(1-d_1)d_2}{d_1d_2}\right)} \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & d_1d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}.$$

类似地做下去，可使

$$\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} \to \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d_1d_2\cdots d_n \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d \end{pmatrix}.$$

其次，根据上述的证明可知

$$A = P_1\cdots P_t\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d \end{pmatrix},$$

其中P_i为消法矩阵，$i=1,2,\cdots,t$。这说明A经若干次行的消法变换得到

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d \end{pmatrix}.$$

据第五章§4命题可知：

$$\det A = \det\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d \end{pmatrix} = d. \quad \text{证完。}$$

定理 3 $m\times n$矩阵A与B相抵的充分必要条件是：存在m阶可

逆阵P与n阶可逆阵Q，使得

$$B=PAQ.$$

证明　必要性．若A与B相抵，则A可经若干次初等变换得到B，也就是

$$P_s\cdots P_1AQ_1\cdots Q_t=B.$$

其中 P_i，Q_j分别为m,n阶的初等矩阵．由于初等矩阵是可逆阵，其积仍是可逆阵，故

$$P=P_s\cdots P_1,\quad Q=Q_1\cdots Q_t$$

是可逆阵，于是

$$B=PAQ.$$

充分性．若存在m阶可逆阵P,n阶可逆阵Q,使得

$$B=PAQ.$$

据本节定理1，有

$$P=P_s\cdots P_1,\quad Q=Q_1\cdots Q_t,$$

其中 P_i,Q_j分别为m,n阶初等矩阵．于是

$$B=P_s\cdots P_1AQ_1\cdots Q_t.$$

据本节命题知，A可经若干次初等变换得到B，故A与B相抵．证完．

推论　**$\text{rank}A=r$充分必要条件是：存在可逆阵P、Q，使得**

$$A=P\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{pmatrix}Q.\quad (r\text{个}1)$$

证明　由于 $\text{rank}A=r$,故A在相抵之下的标准形为

$$\begin{pmatrix}1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0\end{pmatrix}.\quad (r\text{个}1)$$

据定理 3 知，存在可逆阵 P_1、Q_1

$$\begin{pmatrix} 1 & & & & & \\ & \overset{r\text{个}}{\ddots} & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = P_1 A Q_1$$

亦即

$$A = P\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}Q.$$

其中 $P = P_1^{-1}, Q = Q_1^{-1}$.

反之，若

$$A = P\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}Q$$

有

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} = P^{-1}AQ^{-1}.$$

这说明，A在相抵之下的标准形为

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

故 $\operatorname{rank} A = r$. 证完.

最后，讨论乘积阵的行列式与秩.

定理 4 设A、B为任二n阶方阵，则

$$\det(AB)=\det A\cdot\det B.$$

证明 若A是可逆的，则

$$A=P_1P_2\cdots P_s\begin{pmatrix}1&&&\\&\ddots&&\\&&1&\\&&&d\end{pmatrix},\quad \det A=d,$$

其中P_i为消法矩阵，$i=1,2,\cdots,s$. 于是

$$AB=P_1P_2\cdots P_s\begin{pmatrix}1&&&\\&\ddots&&\\&&1&\\&&&d\end{pmatrix}B=P_1P_2\cdots P_sD_n(d)B.$$

据第五章§4命题知，

$$\begin{aligned}\det(AB)&=\det(P_1P_2\cdots P_sD_n(d)B)\\&=\det(D_n(d)B)\\&=d\cdot\det B\\&=\det A\cdot\det B.\end{aligned}$$

若A不是可逆的，这时 $\det A=0$，即 $\operatorname{rank}A=r<n$. 则存在可逆阵P、Q使得

$$A=P\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}Q$$

于是

$$AB=P\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}QB.$$

$$\det(AB)=\det\left(P\cdot\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}QB\right)$$

$$=\det P\cdot\det\left(\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}QB\right)$$

由于

$$\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}QB=\begin{pmatrix}t_{11}&\cdots&t_{1n}\\\cdots&\cdots&\cdots\\t_{r1}&\cdots&t_{rn}\\0&\cdots&0\\\cdots&\cdots&\cdots\\0&\cdots&0\end{pmatrix}.$$

所以

$$\det AB=\det P\cdot\det\begin{pmatrix}t_{11}&\cdots&t_{1n}\\\cdots&\cdots&\cdots\\t_{r1}&\cdots&t_{rn}\\0&\cdots&0\\\cdots&\cdots&\cdots\\0&\cdots&0\end{pmatrix}$$

$$=\det P\cdot 0$$

$$=0.$$

另外，由于 $\det A=0$，得$\det A\cdot\det B=0$。故

$$\det(AB)=\det A\cdot\det B,$$ 证完。

推论 设s个n阶方阵 A_1, A_2, …, A_s, 则

$$\det(A_1A_2\cdots A_s)=\det A_1\cdot\det A_2\cdots\cdots\det A_s.$$

这个推论，只要用一下数学归纳法，便知是正确的。

定理 5 设A、B为二n阶方阵，且A为可逆的，则

$$\operatorname{rank}AB=\operatorname{rank}B,\quad \operatorname{rank}BA=\operatorname{rank}B.$$

证明　据本节定理2的推论，得

$$AB = P_1P_2\cdots P_s\begin{pmatrix}1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d\end{pmatrix}B = P_1P_2\cdots P_sD_n(d)B$$

这说明：B可经若干次行的初等变换得到AB，由于初等变换不改变矩阵的秩，所以有

$$\operatorname{rank}AB = \operatorname{rank}B.$$

类似地可证：$\operatorname{rank}BA = \operatorname{rank}B$。证完。

定理6　设$\boldsymbol{A}$为$\boldsymbol{m}\times\boldsymbol{n}$矩阵，$\boldsymbol{B}$为$\boldsymbol{n}\times\boldsymbol{p}$矩阵，令$\boldsymbol{C}=\boldsymbol{AB}$。若$\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$的秩分别为$r_A$，$r_B$，$r_C$，则

$$r_C \leqslant \min(r_A, r_B).$$

证明　由于$\operatorname{rank}A = r_A$，所以存在可逆阵$P$，$Q$使得

$$A = P\begin{pmatrix}I_{r_A} & 0 \\ 0 & 0\end{pmatrix}Q.$$

于是

$$C = AB = P\begin{pmatrix}I_{r_A} & 0 \\ 0 & 0\end{pmatrix}QB.$$

$$r_c = \operatorname{rank}\left(P\cdot\begin{pmatrix}I_{r_A} & 0 \\ 0 & 0\end{pmatrix}QB\right)$$

$$= \operatorname{rank}\left(\begin{pmatrix}I_{r_A} & 0 \\ 0 & 0\end{pmatrix}QB\right).$$

又因为乘积阵

$$\begin{pmatrix}I_{r_A} & 0 \\ 0 & 0\end{pmatrix}QB$$

里最多有r_A行不为0，所以

$$\operatorname{rank}\left(\begin{pmatrix}I_{r_A} & 0 \\ 0 & 0\end{pmatrix}QB\right) \leqslant r_A.$$

故得

$$r_c \leqslant r_A.$$

同理，由于$\operatorname{rank}B = r_B$，所以存在可逆阵$P_1$，$Q_1$，使得

$$B = P_1\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}Q_1.$$

于是

$$C = AB = AP_1\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}Q_1.$$

$$r_c = \text{rank}\left(AP_1\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}Q_1\right)$$

$$= \text{rank}\, AP_1\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}.$$

又因为乘积阵

$$AP_1\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}$$

里最多有r_B列不为0，所以

$$\text{rank}\left(AP_1\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}\right) \leqslant r_B.$$

故

$$r_c \leqslant r_B$$

总之

$$r_c \leqslant \min(r_A,\ r_B)。\text{证完.}$$

练　习　三

1. 试将下列矩阵A用初等变换化为标准形，并且用初等矩阵将A与其标准形联结成等式：

1) $A = \begin{pmatrix} 0 & 2 & 1 \\ -1 & 1 & 4 \\ 2 & -1 & -3 \end{pmatrix}$，2) $A = \begin{pmatrix} 0 & 0 & 3 & 1 \\ 2 & 1 & -1 & 2 \\ 4 & 2 & 3 & 1 \\ -2 & -1 & 4 & -3 \end{pmatrix}$.

2. 试用初等变换方法求下列矩阵的逆阵：

1) $\begin{pmatrix} 2 & 1 & 7 \\ 5 & 3 & -1 \\ -4 & -3 & 2 \end{pmatrix}$，2) $\begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 0 & 1 & 0 \\ 3 & 1 & 2 & 0 \\ -3 & 1 & 0 & 4 \end{pmatrix}$.

3．试将矩阵

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$$

表成形式为

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} 与 \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$$

的矩阵的乘积。

4．证明：秩为r的矩阵可以表成r个秩为1的矩阵之和。

5．设A是一个n阶方阵，且 $\mathrm{rank}A=r$。证明：存在一个n阶可逆阵P，使得PAP^{-1}的后$n-r$行全为0。

6．设A是一个n阶方阵，且 $|A|=1$。证明：A可以表成消法矩阵的乘积。

7．证明：若$\det A\neq 0$，则$\det A^{-1}=(\det A)^{-1}$。

8．对n阶方阵A，若存在n阶方阵B，使 $AB=I_n$。试问：A是否可逆？B是否可逆？如是可逆的指出它们的逆阵。

9．设n阶方阵A适合 $aA^2+bA+cI_n=0$，$c\neq 0$。证明：A是可逆阵，并求出A^{-1}。

10．设A为$m\times n$矩阵，B为$n\times m$矩阵。证明：当 $m>n$时，方阵$C=AB$的秩小于m。

11．设C为可逆矩阵，试问 ACB 的秩与AB的秩是否一定相等？

12．设A是一个$m\times n$矩阵，B是一个$n\times s$矩阵。证明：

1）若 $\mathrm{rank}A=n$，则 $\mathrm{rank}(AB)=\mathrm{rank}B$；

2）若 $\mathrm{rank}B=n$，则 $\mathrm{rank}(AB)=\mathrm{rank}A$。

习 题 八

1. 设

$$A=\begin{pmatrix}a_1 & 0 & \cdots & 0\\ 0 & a_2 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & a_n\end{pmatrix}.$$

如果A中对角线上的元素 $a_i\neq a_j$ $(i\neq j)$。证明：与A可交换的矩阵只能是对角形阵。

2. 设

$$A=\begin{pmatrix}a_1I_1 & & & \\ & a_2I_2 & & \\ & & \ddots & \\ & & & a_rI_r\end{pmatrix},$$

其中 $a_i\neq a_j$ $(i\neq j)$，I_i是n_i 阶单位阵，$n_1+n_2+\cdots+n_r=n$。证明：与A可交换的矩阵只能是准对角形阵

$$\begin{pmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r\end{pmatrix},$$

其中A_i是n_i阶方阵。

3. 设E_{ij}表示i行j列元素为1，其余元素全为0的n阶方阵，而

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}.$$

证明：

1）若 $AE_{12}=E_{12}A$,则当$k\neq1$时，$a_{k1}=0$；当$k\neq2$时；$a_{2k}=0$；

2）若 $AE_{ij}=E_{ij}A$，则当$k\neq i$时$a_{ki}=0$，当 $k\neq j$时$a_{jk}=0$，且$a_{ii}=a_{jj}$;

3）若A与所有n阶方阵可交换，则A一定是纯量阵，即 aI_n。

4. 设B_1, B_2都与A可交换，则 B_1+B_2, B_1B_2, $B_1^k(k>0)$ 都与A可交换。

5. 设A与B可交换，则$f(A)$与B可交换，其中$f(\lambda)$为任一多项式。

6. 设A, B为n阶方阵。试问当A, B满足什么条件时

$$(A+B)(A-B)=A^2-B^2$$

成立？

7. 设A为n阶方阵。如果对任一$n\times 1$矩阵

$$X=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

都有$AX=0$，那么$A=0$。

8. 试问：

1）两个上（下）三角形阵的乘积是否仍为上（下）三角形阵？

2）可逆的上（下）三角形阵的逆阵是否仍为上（下）三角形阵？

9. 设A是一个n阶方阵，$\operatorname{rank}A=1$。证明：

1）
$$A=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1\ b_2\ \cdots b_n);$$

2）$A^2=kA$，其中k为数。

10. 设A为n阶方阵（$n\geqslant 2$），证明

$$\operatorname{rank}A^*=\begin{cases}n, & 当\ \operatorname{rank}A=n;\\1, & 当\operatorname{rank}A=n-1;\\0, & 当\ \operatorname{rank}A<n-1.\end{cases}$$

11. 设A是n阶方阵($n\geqslant 2$)。证明

$$(A^*)^*=|A|^{n-2}A.$$

12. 设B是r阶方阵，C为$r\times n$阶矩阵，且$\operatorname{rank}C=r$。证明：

1）若 $BC=0$，则 $B=0$；

2）若 $BC=C$，则 $B=I_r$。

13．设A为n阶方阵。如果 $A^2=I_n$，则

$\text{rank}(A+I_n)+\text{rank}(A-I_n)=n$。

14．设A为n阶方阵。如果 $A^2=A$，则

$\text{rank}A+\text{rank}(A-I_n)=n$。

第九章　对称阵在相合之下的标准形

§1　二次型　对称阵在相合之下的分类

在平面解析几何里已经知道，中心在坐标原点的有心二次曲线的化简问题。设已给曲线的方程为

$$ax^2+2bxy+cy^2=d. \qquad (1)$$

于是化简问题就是通过适当的坐标旋转

$$\begin{aligned} x &= x'\cos\theta - y'\sin\theta \\ y &= x'\sin\theta + y'\cos\theta \end{aligned} \qquad (2)$$

把方程（1）的左端

$$f(x,y)=ax^2+2bxy+cy^2 \qquad (3)$$

化简为只含平方项：

$$g(x',y')=a'x'^2+c'y'^2. \qquad (4)$$

类似地，对空间的有心二次曲面也有同样地问题。设中心在原点的二次曲面的方程为

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz=d. \qquad (5)$$

二次曲面（5）的化简问题就是通过适当的坐标旋转，把（5）的左端

$$f(x,y,z)=a_{11}x^2+a_{22}y^2+a_{33}z^2+2a_{12}xy+2a_{13}xz+2a_{23}yz$$

化简为只含平方项：

$$g(x',y',z')=a'_{11}x'^2+a'_{22}y'^2+a'_{33}z'^2.$$

概括地说，（1）与（5）的左端分别是关于两个变数 x,y 和三个变数 x,y,z 的二次齐次函数。而化简问题就是通过适当的变数变

换，使它们对新变数来说只含平方项．这里所用的变数变换的系数构成的矩阵的行列式不等于零，即系数构成的阵为可逆阵．比如，化简（1）所用的变数变换（2）的系数构成的矩阵为

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

其行列式 $det\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \neq 0$。

在许多方面，都要用到 n个（$n \geqslant 3$）变数的二次齐次函数及其化简问题．于是，我们有必要把上述情况加以推广．设关于 n个变数 $x_1, x_2, \cdots, x_n$的二次齐次函数

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{n-1n}x_{n-1}x_n. \qquad (6)$$

叫做二次齐式或叫做二次型．其中 $a_{ij} \in F$（F 为数域）．我们的问题是通过适当的变数变换

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases} \qquad (7)$$

把二次型（6）化简为只含新变数 $y_1, y_2, \cdots, y_n$ 的平方项：

$$f(x_1, x_2, \cdots, x_n) = g(y_1, y_2, \cdots, y_n) = b_1y_1^2 + b_2y_2^2 + \cdots + b_ny_n^2. \qquad (8)$$

其中

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

为可逆方阵，叫做（7）的系数阵．

我们把变数变换（7），当C为可逆阵时，叫做**可逆变数变换**．这时可以说：二次型的问题，就是通过适当的可逆变数变换化为平

方和的形式。

下面利用矩阵乘法把上述的二次型及其化简问题表述出来。

首先，二次型（6）可改写为

$$
\begin{aligned}
f(x_1,x_2,\cdots,x_n)=&a_{11}x_1^2+a_{12}x_1x_2+\cdots+a_{1n}x_1x_n\\
&+a_{21}x_2x_1+a_{22}x_2^2+\cdots+a_{2n}x_2x_n\\
&+\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\
&+a_{n1}x_nx_1+a_{n2}x_nx_2+\cdots+a_{nn}x_n^2
\end{aligned}
\tag{9}
$$

其中 $a_{ij}=a_{ji}$，$i,j=1,2,\cdots,n$。

于是，又可写为

$$
f(x_1,x_2,\cdots,x_n)=(x_1x_2\cdots x_n)\begin{pmatrix}a_{11}a_{12}\cdots a_{1n}\\ a_{21}a_{22}\cdots a_{2n}\\ \cdots\cdots\cdots\cdots\\ a_{n1}a_{n2}\cdots a_{nn}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}
$$

或 $$f(x_1,x_2,\cdots,x_n)=(x_1x_2\cdots x_n)A\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}\tag{10}$$

或 $$f(x_1,x_2,\cdots,x_n)=X'AX.\tag{11}$$

其中

$$
A=\begin{pmatrix}a_{11}a_{12}\cdots a_{1n}\\ a_{21}a_{22}\cdots a_{2n}\\ \cdots\cdots\cdots\cdots\\ a_{n1}a_{n2}\cdots a_{nn}\end{pmatrix},\quad X=\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}.
$$

A是二次型（6）的系数构成的 n阶方阵，其主对角线上的元素依次是平方项 $x_1^2,x_2^2,\cdots$，x_n^2 的系数 $a_{11},a_{22},\cdots,a_{nn}$；主对角线右上方的元素 $a_{ij}(i<j)$ 就是 x_ix_j 的系数 $2a_{ij}$ 的一半 a_{ij}，其余的一半正好按主对角线对称地记在 (j,i)位置，即 $a_{ij}=a_{ji}$。

这样 n阶矩阵 $A=(a_{ij})$有性质：$A'=A$，把有这样性质的方阵叫做对称阵。

我们称(10)或(11)为二次型 $f(x_1, x_2, \cdots, x_n)$ 的矩阵形式，A叫做$f(x_1, x_2, \cdots, x_n)$ 的表示矩阵。于是任一二次型都有唯一确定的矩阵形式和一个 n阶对称阵做为表示矩阵。反之，任一n阶对称阵 A 都是某一确定的 n个变数的二次型的表示矩阵。换句话说，在数域 F 上，一切 n 个变数的二次型与全体 n阶对称阵之间建立了一对一的对应关系：①

$$f(x_1, x_2, \cdots, x_n) = (x_1, x_2, \cdots, x_n)\ A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \longleftrightarrow A, \text{其中} A' = A.$$

其次看用变数变换（7）化简二次型（6）的问题。先把(7)写成矩阵形式：

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = C\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{12}$$

或 $$X = CY, \text{其中} X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

于是把（12）代入（10），则有

$$f(x_1, x_2, \cdots, x_n) = \left(C\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right)' A \cdot \left(C\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right)$$

---

①这里需要说明一点：二次型$f(x_1, x_2, \cdots, x_n)$与 $g(x_1, x_2, \cdots, x_n)$相同必要而且只要其表示矩阵相同。

$$= (y_1 y_2 \cdots y_n) C'AC \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$= g(y_1, y_2, \cdots, y_n) \tag{13}$$

易知 $C'AC$ 为对称阵。

要求 $g(y_1, y_2, \cdots, y_n)$ 只含 $y_1, y_2, \cdots, y_n$ 的平方项必要而且只要 $C'AC$ 是一个对角形阵：

$$C'AC = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \cdots & \\ & & & d_n \end{pmatrix}.$$

这样便有

$$g(y_1, y_2, \cdots, y_n) = (y_1 y_2 \cdots y_n) \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \cdots & \\ & & & d_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$= d_1 y_1^2 + d_2 y_2^2 + \cdots + d_n y_n^2.$$

因此化简二次型（6）的问题，就转化成如何求出一个可逆阵 C，使 $C'AC$ 为对角阵？即用可逆阵 C，按

$$A \longrightarrow C'AC$$

的方式化简对称阵 A 成对角形阵的问题。

最后我们指出，这正是关于对称阵的一种分类和标准形问题。为此我们自然地引入

定义 设 A、B 为任二 n 阶方阵，如果有可逆阵 P，使

$$B = P'AP,$$

则说 A 与 B 相合，记作 $A \simeq B$。这时称 P 为 A 相合于 B 的演化阵，简称 P 为相合演化阵。

命题 1 n 阶方阵的相合关系具有下述三条性质：

1）反身性 $A \simeq A$；

2）**对称性**　若 $A\simeq B$，则 $B\simeq A$；

3）**传递性**　若 $A\simeq B$，$B\simeq C$，则 $A\simeq C$。

事实上，由 $I'AI=A$，有 $A\simeq A$；

若 $A\simeq B$，即有可逆阵 P，使 $B=P'AP$。于是，$A=P'^{-1}BP^{-1}=(P^{-1})'BP^{-1}$，故 $B\simeq A$；

若 $A\simeq B$，$B\simeq C$。即有可逆阵 P,Q 使 $B=P'AP$，$C=Q'BQ$。于是 $C=Q'P'APQ=(PQ)'A(PQ)$，故 $A\simeq C$。本命题得证。

命题2　A为对称阵的充分必要条件是：$P'AP$ 为对称阵，其中P 为可逆阵。

证明　必要性。因为

$$(P'AP)'=P'A'(P')'=P'AP,$$

故 $P'AP$为对称阵。

充分性　因 $P'AP$为对称阵，所以一方面

$$(P'AP)'=P'AP,$$

另一方面

$$(P'AP)'=P'A'(P')'=P'A'P$$

于是，有

$$P'A'P=P'AP.$$

把上式两端从左边乘P'^{-1}，从右边乘 P^{-1}，得

$$A'=A.$$　证完。

命题1表明，n 阶方阵的相合关系能够把一切 n 阶方阵区分成类；

A与B分在一类必要而且只要A与B相合。

命题2表明，把相合关系仅限于用在对称阵，恰能把一切 n 阶对称阵区分成类。

这样一来，用可逆阵C，按

$$A\longrightarrow C'AC$$

的方式化简对称阵A成对角形阵的问题，这正是一切 n 阶对称阵在相合分类之下，于每一类中求出对角形阵为其标准形的问题。

以下主要讨论实对称阵（矩阵的元素取实数）在相合之下的标准形问题。在此基础上容易指出关于复数域上相应的结论。因此，今后凡是说到矩阵都是指实矩阵。

再有，实对称阵在相合之下的化简问题本是实二次型（系数取实数）在可逆的变数变换之下的化简问题的一种表现形式。为了简便，我们不把对称阵的每一结果都用相应的二次型形式表现出来，只是在适当的地方写出它们的二次型的相应结论。

练 习 一

1．试写出以下列对称阵为表示矩阵的二次型：

$$A=\begin{pmatrix}1 & 2 & -\frac{3}{2}\\ 2 & 0 & 1\\ -\frac{3}{2} & 1 & -1\end{pmatrix}\qquad B=\begin{pmatrix}0 & 1 & 2 & -1\\ 1 & 0 & 3 & 1\\ 2 & 3 & 0 & -\frac{1}{2}\\ -1 & 1 & -\frac{1}{2} & 0\end{pmatrix}.$$

2．试写出下列二次型的表示矩阵，且写成矩阵形式。

1) $f(x_1,x_2,x_3)=2x_1^2-3x_1x_2+2x_1x_3+2x_2^2-5x_2x_3$；

2) $f(x_1,x_2,x_3,x_4)=3x_1x_2+4x_1x_3+5x_1x_4-2x_2x_3-2x_2x_4$
$-3x_3x_4$。

3．证明：若 A、B均为对称阵，则 $A+B$，kA都是对称阵。

4．试给出秩为2、3、4的对称阵，而它们的阶数分别为2、4、7。

§2 对称阵在成套的初等变换下的化简

我们先用初等变换的观点来分析相合的两个对称阵之间的关系。

设 A、B 为两个 n 阶对称阵，它们是相合的。即存在可逆阵 P，使

$$P'AP=B.$$

因为可逆阵都能表成一些初等矩阵的乘积，所以可令

$$P=P_1P_2\cdots P_s,$$

其中 P_i皆为初等阵， $i=1,2,\cdots, s$。于是

$$P'=P_s'\cdots P_2'P_1',$$

从而，有

$$P_s'\cdots P_2'P_1'AP_1P_2\cdots P_s=B.$$

我们来具体考察 $P_1'AP_1$ 的结果.

1）若 $P_1=D_i(k)$，即 P_1 是倍法阵，于是 $P_1'=D_i'(k)=D_i(k)$。故有

$$P_1'AP_1=D_i(k)AD_i(k).$$

这结果对A做两次初等变换，即用 $k(\neq 0)$ 乘A的第 i 行，接着再用 k乘 A 的第 i 列。

2）若 $P_1=P_{ij}(k)$，即 P_1 是消法阵，为了具体不妨令 $P_1=P_{12}(k)$。于是 $P_{12}'(k)=P_{21}(k)$，故

$$P_1'AP_1=P_{21}(k)AP_{12}(k).$$

这结果也是对A做两次初等变换：把第一行的 k 倍加于第二行，接着再把第一列的 k倍加于第二列.

3）若$P_1=C_{ij}$，即 P_1 是换法阵．于是 $C'_{ij}=C_{ij}$，故有

$$P_1'AP_1=C_{ij}AC_{ij}.$$

这结果对A还是做两次初等变换：交换A的第 i 行与第 j 行，接着再交换第 i 列与第 j 列.

象上边那样，对阵A同时做两次初等变换，叫做对A做一套初等变换，并且对对称阵A每做完一套初等变换得到的新阵仍为对称阵。于是

$$P_2'(P_1'AP_1)P_2$$

就是对A做两套初等变换。一般的有

$$P_s'\cdots P_2'P_1'AP_1P_2\cdots P_s$$

就是对A做了 s 套初等变换．这样，

$$P'_s\cdots P'_2P'_1AP_1P_2\cdots P_s=B$$

表明A经 s 套初等变换变为 B。因此，两个阵A与 B 相合又可理解为：

A与 B 相合必要而且只要经若干套初等变换把A变到了 B。按此理解，对称阵A在相合分类下的标准形的问题就可以说成：用成套的初等变换把对称阵A化成最简形式。有了这样的理解和说法，不难得到对称阵在相合分类之下的标准形。

命题 1　设对称阵

$$A=\begin{pmatrix} a_{11}a_{12}\cdots a_{1n} \\ a_{21}a_{22}\cdots a_{2n} \\ \cdots\cdots\cdots\cdots \\ a_{n1}a_{n2}\cdots a_{nn} \end{pmatrix}$$

的(1,1)位置 $a_{11}\neq 0$，那么用若干次成套的初等变换可以把A变为

$$B=\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

其中 A_1 为 $n-1$ 阶对称阵。

证明　如果 $a_{21}\neq 0$，则用一套消法变换 $P_{21}\left(-\frac{a_{21}}{a_{11}}\right)$，$P_{12}\left(-\frac{a_{12}}{a_{11}}\right)$变$A$，就得

$$P'_{12}\left(-\frac{a_{12}}{a_{11}}\right)AP_{12}\left(-\frac{a_{12}}{a_{11}}\right)=P_{21}\left(-\frac{a_{21}}{a_{11}}\right)AP_{12}\left(-\frac{a_{12}}{a_{11}}\right)$$

$$=\begin{pmatrix} a_{11} & 0 & a_{13} & \cdots & a_{1n} \\ 0 & & & & \\ a_{31} & & & & \\ \vdots & & * & & \\ a_{n1} & & & & \end{pmatrix}=B_1.$$

类似地，若还有 $a_{31}\neq 0$，则用一套消法变换$P_{31}\left(-\frac{a_{31}}{a_{11}}\right)$，$P_{13}\left(-\frac{a_{13}}{a_{11}}\right)$变$B_1$，就得

$$P'_{13}\left(-\frac{a_{13}}{a_{11}}\right)P'_{12}\left(-\frac{a_{12}}{a_{11}}\right)AP_{12}\left(-\frac{a_{12}}{a_{11}}\right)P_{13}\left(-\frac{a_{13}}{a_{11}}\right)$$

$$=\begin{pmatrix} a_{11} & 0 & 0 & a_{14} & \cdots & a_{1n} \\ 0 & & & & & \\ 0 & & & & & \\ a_{41} & & & * & & \\ \vdots & & & & & \\ a_{n1} & & & & & \end{pmatrix}.$$

把上述手续如此继续做下去，至多用 $n-1$ 次成套的消法变换，把A变为：

$$B=\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

因为变得的 B 是对称阵，故 A_1 为 $n-1$ 阶对称阵。

命题 2　若A为非零对称阵，那么用成套的初等变换必能使(1,1)位置的元素不等于零。

证明　我们分两种情形考虑：

1）A的对角线上有非零元素 $a_{ii}\neq 0$。这时对A做两次换法变换就把 a_{ii} 换到（1，1）位置上：

$$C_{1i}AC_{1i}=\begin{pmatrix} a_{ii} & \cdots \\ \vdots & * \end{pmatrix}.$$

2）A的对角线上元素全为 0，但有 $a_{ij}\neq 0\ (i<j)$。这时对A做两次消法变换，就能使 (i,i) 位置变为 $2a_{ij}$：

$$P_{ij}(1)AP_{ji}(1)=\begin{pmatrix} \ddots & \vdots & * \\ \cdots & 2a_{ij} & \cdots \\ * & \vdots & \ddots \end{pmatrix}\ (i\text{行})。$$

$(i\text{ 列})$

这样就又归结为1）了。

定理1　任意对称阵A都能用成套的初等变换化成对角形阵。换句话说，任意对称阵A都相合于对角形阵。

证明　当$A=0$时，A已是对角形阵。以下假设$A\neq 0$，依命题2可认为A的（1，1）位置的元素$a_{11}=\lambda_1\neq 0$。再由命题1，A相合于

$$B=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

这里A_1是$n-1$阶对称阵。

对A_1重复以上的论述（或对阵A的阶数n用归纳法），便可推出A相合于对角形阵：

$$\begin{pmatrix}\lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \lambda_r & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0\end{pmatrix}, \tag{1}$$

其中$\lambda_1,\lambda_2,\cdots,\lambda_r$均不为0。证完。

对角形阵（1）中对角线上非零元素的个数r就是对称阵A的秩。显然，r是相合分类的不变量。但它不是完备不变量，即秩数相同的对称阵A, B未必相合。比如

$$A=\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix}\text{ 与 }B=\begin{pmatrix}-1 & & \\ & -1 & \\ & & -1\end{pmatrix}$$

的秩都是3，可是A与B不相合。

为了找出对称阵在相合分类之下的完备不变量，从而完全确定对称阵的相合分类，必须进一步化简对角阵（1），从中发现相合对称阵的基本共性。

定理2 任意对称阵A都能相合于这样特殊的对角形阵：

$$\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}, \tag{2}$$

其中 $p+q=\text{rank}A$，称（2）为A的标准形。

证明　由定理1知A相合于对角形阵：

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \lambda_2 & & & & & \\ & & \ddots & & & & \\ & & & \lambda_r & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}.$$

因每个 λ_i 皆不等于零，故可设当中有p个正数，$r-p=q$ 个负数。于是用适当的成套的换法变换，可以把这个对角形阵变成另一个与其相合的对角形阵：

$$\begin{pmatrix} a_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & a_p & & & & & & \\ & & & -b_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -b_q & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix} \tag{3}$$

其中 $a_i>0(i=1,\cdots,p)$；$b_j>0(j=1,\cdots,q)$。

再对（3）做一些成套的初等变换，则得

$$D_r\left(\frac{1}{\sqrt{b_q}}\right)\cdots D_{p+1}\left(\frac{1}{\sqrt{b_1}}\right)D_p\left(\frac{1}{\sqrt{a_p}}\right)\cdots D_1\left(\frac{1}{\sqrt{a_1}}\right)\times$$

$$\begin{pmatrix} a_1 & & & & & & & \\ & \ddots & & & & & & \\ & & a_p & & & & & \\ & & & -b_1 & & & & \\ & & & & \ddots & & & \\ & & & & & -b_q & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix} D_1\left(\frac{1}{\sqrt{a_1}}\right)\cdots$$

$$D_p\left(\frac{1}{\sqrt{a_p}}\right)D_{p+1}\left(\frac{1}{\sqrt{b_1}}\right)\cdots D_r\left(\frac{1}{\sqrt{b_q}}\right)$$

$$=\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}=\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}.$$ 证完。

把定理 1 与定理 2 用二次型的语言表达出来，便有

定理1′ 对任意实系数二次型

$$f(x_1,\cdots,x_n)=(x_1\cdots x_n)A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

都有适当的变数变换

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}=Q\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \det Q\neq 0$$

使 $f(x_1,\cdots,x_n)$ 化为平方和：

$$g(y_1,\cdots,y_n)=\lambda_1 y_1^2+\cdots+\lambda_r y_r^2 \tag{4}$$

定理2′ 对任意实二次型

$$f(x_1,\cdots,x_n)=(x_1\cdots x_n)A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

都有适当的变数变换

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \qquad \det P \neq 0,$$

使

$$f(x_1, \cdots, x_n) = g(y_1, \cdots, y_n) = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2, \tag{5}$$

称（5）为二次型 $f(x_1, \cdots, x_n)$ 的标准形。

定理1和定理2的证明过程就是化简对称阵成对角形阵（1）与（2）的具体方法。而且能同时得到体现所用的那些成套初等变换的Q、P。这是因为，当用成套的初等变换，把对称阵化为（1）或（2）时，即

$$Q_t' \cdots Q_1' A Q_1 \cdots Q_t = \begin{pmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \lambda_r & & \\ & & & & 0 & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

或

$$P_s' \cdots P_1' A P_1 \cdots P_s = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & 0 & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}$$

时，而

$$I_n Q_1 \cdots Q_t \text{ 与 } I_n P_1 \cdots P_s$$

就是所求的Q与P。就是说，对A施行成套初等变换的同时，对单位阵I_n施行同样列的变换，当把A化为（1）或（2）时，单位阵I_n也就化为Q或P了。

例 1 用成套的初等变换化简对称阵 A 成对角形阵，其中

$$A=\begin{pmatrix}-1 & 2 & 0 & -2\\ 2 & -3 & 0 & 1\\ 0 & 0 & 0 & 0\\ -2 & 1 & 0 & 2\end{pmatrix}.$$

解 把 A 写出后，再把单位阵 I_n 写在 A 的下面，这样便于一起进行列的初等变换.

$$\begin{pmatrix}-1 & 2 & 0 & -2\\ 2 & -3 & 0 & 1\\ 0 & 0 & 0 & 0\\ -2 & 1 & 0 & 2\\ \cdots & \cdots & \cdots & \cdots\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{matrix}\text{上面是 } A\\ \\ \text{下面是 } I_4\end{matrix}.$$

把第一行的 2 倍、－2 倍分别加于第二行、第四行；接着把第一列的 2 倍、－2 倍分别加于第二列、第四列（这时实际上把第一行与第一列的 2 、－2 都化为 0 了），则得

$$\begin{pmatrix}-1 & 0 & 0 & 0\\ 0 & 1 & 0 & -3\\ 0 & 0 & 0 & 0\\ 0 & -3 & 0 & 6\\ 1 & 2 & 0 & -2\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

把第二行的 3 倍加于第四行，然后把二列的 3 倍加于第四列，则得

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \\ 1 & 2 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (6)$$

到此A已化为对角形阵了，共用了三套初等变换，把它们积累起来就是可逆阵

$$Q=\begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

即

$$Q'AQ=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 4 & 3 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 2 & 0 & -2 \\ 2 & -3 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$=\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}.$$

进而在此基础上可以把A化简成更简单的对角形，即标准形。

用$\frac{1}{\sqrt{3}}$乘（6）的第四行，接着用$\frac{1}{\sqrt{3}}$乘（6）的第四列，则得

$$\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 2 & 0 & \frac{4}{\sqrt{3}} \\ 0 & 1 & 0 & \frac{3}{\sqrt{3}} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

最后交换一、二行，接着交换一、二列；再交换三、四行，接着交换三、四列，则得

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 1 & \frac{4}{\sqrt{3}} & 0 \\ 1 & 0 & \frac{3}{\sqrt{3}} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 \end{pmatrix}.$$

这时把从A化到标准形所用的列的初等变换积累起来就得到可逆阵

$$P=\begin{pmatrix} 2 & 1 & \frac{4}{\sqrt{3}} & 0 \\ 1 & 0 & \frac{3}{\sqrt{3}} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 \end{pmatrix},$$

而A的标准形为

$$P'AP=\begin{pmatrix}1&&&\\&-1&&\\&&-1&\\&&&0\end{pmatrix}$$

例 2　求对称阵

$$A=\begin{pmatrix}1&-1&1&-1\\-1&1&1&-2\\1&1&1&0\\-1&-2&0&-2\end{pmatrix}$$

的标准形及相合演化阵P。

解

$$\begin{pmatrix}1&-1&1&-1\\-1&1&1&-2\\1&1&1&0\\-1&-2&0&-2\\1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\xrightarrow[P_{12}(1),P_{13}(-1),P_{14}(1)]{P_{21}(1),P_{31}(-1),P_{41}(1)}$$

$$\begin{pmatrix}1&0&0&0\\0&0&2&-3\\0&2&0&1\\0&-3&1&-3\\1&1&-1&1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}\xrightarrow[P_{32}(1)]{P_{23}(1)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 2 & -2 \\ 0 & 2 & 0 & 1 \\ 0 & -2 & 1 & -3 \\ 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{23}\left(-\frac{1}{2}\right),\ P_{24}\left(\frac{1}{2}\right)]{P_{32}\left(-\frac{1}{2}\right),\ P_{42}\left(\frac{1}{2}\right)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & -1 & 2 \\ 0 & 0 & 2 & -4 \\ 1 & 1 & -\frac{3}{2} & \frac{3}{2} \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{34}(2)]{P_{43}(2)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & -\frac{3}{2} & -\frac{3}{2} \\ 0 & 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[D_2\left(\frac{1}{2}\right)]{D_2\left(\frac{1}{2}\right)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & \frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

A的标准形为

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 0 \end{pmatrix},$$

而相合演化阵

$$P = \begin{pmatrix} 1 & \frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

做为对称阵标准形的应用，来化简二次型。

例 3　化二次型

$$f(x_1,x_2,x_3)=2x_1x_2+2x_1x_3-6x_2x_3$$

为标准形并求出其可逆变数变换.

解 先写出二次型 $f(x_1,x_2,x_3)$ 的表示矩阵

$$A=\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\end{pmatrix}.$$

其次用成套的初等变换化表示矩阵 A 为标准形:

$$\begin{pmatrix}0&1&1\\1&0&-3\\1&-3&0\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}\xrightarrow[P_{21}(1)]{P_{12}(1)}\begin{pmatrix}2&1&-2\\1&0&-3\\-2&-3&0\\1&0&0\\1&1&0\\0&0&1\end{pmatrix}\xrightarrow[P_{12}\left(-\frac{1}{2}\right),\ P_{13}(1)]{P_{21}\left(-\frac{1}{2}\right),\ P_{31}(1)}$$

$$\begin{pmatrix}2&0&0\\0&-1&-2\\0&-2&-4\\1&-\frac{1}{2}&1\\1&\frac{1}{2}&1\\0&0&1\end{pmatrix}\xrightarrow[P_{23}(-2)]{P_{32}(-2)}\begin{pmatrix}2&0&0\\0&-1&0\\0&0&4\\1&-\frac{1}{2}&2\\1&\frac{1}{2}&0\\0&0&1\end{pmatrix}$$

这时已把表示矩阵 A 化为对角形了，相应的把二次型化为平方和的形式了:

$$f(x_1,x_2,x_3)=g(y_1,y_2,y_3)=2y_1^2-y_2^2+4y_3^2$$

其可逆变数变换为

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1&-\frac{1}{2}&2\\1&\frac{1}{2}&0\\0&0&1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}.$$

进一步化标准形:

$$
\begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \\ 1 & -\frac{1}{2} & 2 \\ 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}
\xrightarrow[D_1\left(\frac{1}{\sqrt{2}}\right),\ D_3\left(\frac{1}{2}\right)]{D_1\left(\frac{1}{\sqrt{2}}\right),\ D_3\left(\frac{1}{2}\right)}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}
$$

$$
\xrightarrow[C_{23}]{C_{23}}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \\ \frac{1}{\sqrt{2}} & 1 & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.
$$

于是二次型的表示矩阵已化为标准形，那么二次型的标准形为

$$
f(x_1, x_2, x_3) = h(z_1, z_2, z_3) = z_1^2 + z_2^2 - z_3^2,
$$

其可逆变数变换为

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 1 & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.
$$

上述化对称阵为标准形、化二次型为标准形都是用成套的初等变换化的，所以我们称为**初等变换法**。

下面再介绍一种方法，叫做**配方法**。其实质也是初等变换，并且是成套的初等变换，只不过是这里每做一步是若干次成套初等变换的积累罢了。比如，例 2 开头对 A 做三套初等变换：

$P_{41}(1)P_{31}(-1)P_{21}(1)AP_{12}(1)P_{13}(-1)P_{14}(1)$

$$
=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 2 & 0 & 1 \\ 0 & -3 & 1 & -3 \end{pmatrix}.
$$

而

$$
P=P_{12}(1)P_{13}(-1)P_{14}(1)=\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

$$
\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

那么

$$
P'AP=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -2 \\ 1 & 1 & 1 & 0 \\ -1 & -2 & 0 & -2 \end{pmatrix}
$$

$$
\begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & -3 \\ 0 & 2 & 0 & 1 \\ 0 & -3 & 1 & -3 \end{pmatrix}
$$

这说明接着分别做三次成套的初等变换，与把这三次成套的初等变换积累起来一次去做，效果完全一样，都是把阵A的第一行、第一列除(1,1)位置外其余元素均化为0。

现在就来介绍用配方法化对称阵为标准形。

设对称阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad (a_{ij}=a_{ji}).$$

如$A=0$，已是标准形了。令$A\neq 0$，分两种情况：

1）若A中对角线上至少有一元素不为0，不妨假设 $a_{11}\neq 0$。于是用可逆阵

$$R=\begin{pmatrix} 1 & -a_{11}^{-1}a_{12} & \cdots & -a_{11}^{-1}a_{1n} \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{7}$$

来化A，即

$$R'AR=\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -a_{11}^{-1}a_{12} & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ -a_{11}^{-1}a_{1n} & 0 & \cdots & 1 \end{pmatrix}\begin{pmatrix} a_{11}a_{12}\cdots a_{1n} \\ a_{21}a_{22}\cdots a_{2n} \\ \cdots\cdots\cdots \\ a_{n1}a_{n2}\cdots a_{nn} \end{pmatrix}$$

$$\begin{pmatrix} 1 & -a_{11}^{-1}a_{12} & \cdots & -a_{11}^{-1}a_{1n} \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{pmatrix}\begin{pmatrix} 1 & -a_{11}^{-1}a_{12} & \cdots & -a_{11}^{-1}a_{1n} \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$=\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a'_{22} & \cdots & a'_{2n} \\ \cdot & \cdots & A_1 & \cdots \\ 0 & a'_{n2} & \cdots & a'_{nn} \end{pmatrix}.$$

这样把阵A化为与其相合的，并且除（1,1）位置是 a_{11} 外，第一行及第一列其余元素都是0，而A_1仍为对称阵。

2）若A中对角线上的元素均为0，而有某 $a_{ij}\neq 0$，不妨假定$a_{12}\neq 0$。于是用可逆阵

$$S=\begin{pmatrix}1 & 1 & 0 & \cdots & 0\\ 1 & -1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix} \tag{8}$$

来化A。即

$$S'AS=\begin{pmatrix}1 & 1 & 0 & \cdots & 0\\ 1 & -1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}\begin{pmatrix}0 & a_{12} & \cdots & a_{1n}\\ a_{21} & 0 & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & 0\end{pmatrix}\begin{pmatrix}1 & 1 & 0 & \cdots & 0\\ 1 & -1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}$$

$$=\begin{pmatrix}a_{21} & a_{12} & a'_{13} & \cdots & a'_{1n}\\ -a_{21} & a_{12} & a'_{23} & \cdots & a'_{2n}\\ a_{31} & a_{32} & 0 & \cdots & a_{3n}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & a_{n3} & \cdots & 0\end{pmatrix}\begin{pmatrix}1 & 1 & 0 & \cdots & 0\\ 1 & -1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 1\end{pmatrix}$$

$$=\begin{pmatrix}2a_{12} & 0 & a'_{13} & \cdots & a'_{1n}\\ 0 & -2a_{12} & a'_{23} & \cdots & a'_{2n}\\ a'_{31} & a'_{32} & 0 & \cdots & a'_{3n}\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ a'_{n1} & a'_{n2} & a'_{n3} & \cdots & 0\end{pmatrix}$$

这就是说，当A的对角线元素全为0时，可用阵S来化A，使得到与A相合的阵 $S'AS$ 中（1,1）位置元素不为0.

综上所述，配方法化对称阵A为标准形的方法是：

若A的对角线元素不全为0，我们可以假定 $a_{11}\neq 0$，利用（7）型矩阵R化A，即

$$A\longrightarrow R'AR=B.$$

这时B中除（1,1）位置外，第一行、第一列其余元素均为0；

若A的对角线上元素全为0，则可用（8）型矩阵S来化A，即

$$A \longrightarrow S'AS = B_1$$

这时B_1的对角线上元素就不全为0了。然后接着用（7）型阵R来化。这样反复用（7）型阵R或（8）型阵S就可将A化为对角形阵，然后用成套的倍法变换的积累起来的阵来化（有时需要适当的交换行、列）就可以成为标准形了。与此同时把每步相合演化阵从左至右依次乘起来就是所求的把A化为标准形的相合演化阵。

例4　将对称阵

$$A=\begin{pmatrix} 0 & 2 & 1 & -3 \\ 2 & 0 & -2 & 1 \\ 1 & -2 & 0 & 1 \\ -3 & 1 & 1 & 0 \end{pmatrix}$$

化为标准形，并求出相合演化阵。

解　因为阵A对角线上元素全为0，故利用（8）型阵S_1来化A，即

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 2 & 1 & -3 \\ 2 & 0 & -2 & 1 \\ 1 & -2 & 0 & 1 \\ -3 & 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$=\begin{pmatrix} 4 & 0 & -1 & -2 \\ 0 & -4 & 3 & -4 \\ -1 & 3 & 0 & 1 \\ -2 & -4 & 1 & 0 \end{pmatrix}.$$

接着用（7）型阵R_1来化，即

$$
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{4} & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 4 & 0 & -1 & -2 \\ 0 & -4 & 3 & -4 \\ -1 & 3 & 0 & 1 \\ -2 & -4 & 1 & 0 \end{pmatrix}
\begin{pmatrix} 1 & 0 & \frac{1}{4} & \frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

$$
=\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & -4 & 3 & -4 \\ 0 & 3 & -\frac{1}{4} & \frac{1}{2} \\ 0 & -4 & \frac{1}{2} & -1 \end{pmatrix}.
$$

再用（7）型阵 R_2来化，即

$$
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{3}{4} & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & -4 & 3 & -4 \\ 0 & 3 & -\frac{1}{4} & \frac{1}{2} \\ 0 & -4 & \frac{1}{2} & -1 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{3}{4} & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

$$
=\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & \frac{5}{4} & -\frac{5}{2} \\ 0 & 0 & -\frac{5}{2} & 3 \end{pmatrix}
$$

继续用（7）型阵 R_3 来化，即

$$
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix}
\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & \frac{5}{4} & -\frac{5}{2} \\ 0 & 0 & -\frac{5}{2} & 3 \end{pmatrix}
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}
$$

$$
=\begin{pmatrix} 4 & & & \\ & -4 & & \\ & & \frac{5}{4} & \\ & & & -2 \end{pmatrix}.
$$

现在已化A为对角形了，进一步用

$$
D=\begin{pmatrix} \frac{1}{2} & & & \\ & \frac{1}{2} & & \\ & & \frac{2}{\sqrt{5}} & \\ & & & \frac{1}{\sqrt{2}} \end{pmatrix}
$$

来化，即

$$
\begin{pmatrix} \frac{1}{2} & & & \\ & \frac{1}{2} & & \\ & & \frac{2}{\sqrt{5}} & \\ & & & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 4 & & & \\ & -4 & & \\ & & \frac{5}{4} & \\ & & & -2 \end{pmatrix}\begin{pmatrix} \frac{1}{2} & & & \\ & \frac{1}{2} & & \\ & & \frac{2}{\sqrt{5}} & \\ & & & \frac{1}{\sqrt{2}} \end{pmatrix}
$$

$$
=\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}.
$$

最后用

$$
C_{23}=\begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{pmatrix}.
$$

来化，即

$$
\begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{pmatrix}\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}\begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{pmatrix}
$$

$$
=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}.
$$

故

$$\begin{pmatrix}1&&&\\&1&&\\&&-1&\\&&&-1\end{pmatrix}$$

为A的标准形，其相合演化阵

$$P=S_1R_1R_2R_3DC_{23}=\begin{pmatrix}\frac{1}{2}&\frac{2}{\sqrt{5}}&\frac{1}{2}&\frac{3}{2\sqrt{2}}\\ \frac{1}{2}&-\frac{1}{\sqrt{5}}&-\frac{1}{2}&\frac{1}{2\sqrt{2}}\\ 0&\frac{2}{\sqrt{5}}&0&\frac{2}{\sqrt{2}}\\ 0&0&0&\frac{1}{\sqrt{2}}\end{pmatrix}.$$

这里已经看得很清楚，（7）型阵R是若干个消法阵之积，而（8）型阵是消法阵与倍法阵之积，故配方法实质就是初等变换法。那么为什么叫配方法呢？这个名称来源于二次型化标准形。我们用例子来说明。

例5　将二次型

$$f(x_1,x_2,x_3)=2x_1^2-4x_1x_2+6x_1x_3+3x_2x_3$$

化为标准形，并求出其可逆变数变换。

解　因二次型有平方项 $2x_1^2$，所以按 x_1 来并项配方，

$$\begin{aligned}f(x_1,x_2,x_3)&=2x_1^2-4x_1x_2+6x_1x_3+3x_2x_3\\&=2(x_1^2-2x_1x_2+3x_1x_3)+3x_2x_3\\&=2[x_1^2-2x_1(x_2-\frac{3}{2}x_3)+(x_2-\frac{3}{2}x_3)^2]\\&\quad-2(x_2-\frac{3}{2}x_3)^2+3x_2x_3\\&=2[x_1-(x_2-\frac{3}{2}x_3)]^2-2(x_2^2-3x_2x_3+\frac{9}{4}x_3^2)\\&\quad+3x_2x_3\\&=2(x_1-x_2+\frac{3}{2}x_3)^2-2x_2^2+9x_2x_3-\frac{9}{2}x_3^2.\end{aligned}$$

令变数变换

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & \frac{3}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \text{ 或} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -\frac{3}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \tag{9}$$

代入上式，得

$$f(x_1, x_2, x_3) = g(y_1, y_2, y_3) = 2y_1^2 - 2y_2^2 + 9y_2y_3 - \frac{9}{2}y_3^2. \tag{10}$$

接着按 y_2 并项配方：

$$\begin{aligned} g(y_1, y_2, y_3) &= 2y_1^2 - 2\left(y_2^2 - \frac{9}{2}y_2y_3\right) - \frac{9}{2}y_3^2 \\ &= 2y_1^2 - 2\left[y_2^2 - 2y_2 \cdot \frac{9}{4}y_3 + \left(\frac{9}{4}y_3\right)^2\right] + \frac{81}{8}y_3^2 - \frac{9}{2}y_3^2 \\ &= 2y_1^2 - 2\left(y_2 - \frac{9}{4}y_3\right)^2 + \frac{45}{8}y_3^2 \end{aligned}$$

令变数变换

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{9}{4} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \text{ 或} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{9}{4} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \tag{11}$$

代入上式，得

$$h(z_1, z_2, z_3) = 2z_1^2 - 2z_2 + \frac{45}{8}z_3^2. \tag{12}$$

(12) 式已是平方和的形式．进一步令变数变换

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & & \\ & \frac{1}{\sqrt{2}} & \\ & & \frac{\sqrt{8}}{\sqrt{45}} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \tag{13}$$

代入上式，得

$$k(u_1, u_2, u_3) = u_1^2 - u_2^2 + u_3^2.$$

再利用变数变换

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 0 & 1 \\ & 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \tag{14}$$

代入上式，得

$$l(v_1, v_2, v_3) = v_1^2 + v_2^2 - v_3^2. \tag{15}$$

(15) 就是二次型 $f(x_1, x_2, x_3)$ 的标准形，其把 $f(x_1, x_2, x_3)$ 化为 (15) 的变数变换为：

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -\frac{3}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{9}{4} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & & \\ & \frac{1}{\sqrt{2}} & \\ & & \frac{\sqrt{8}}{\sqrt{45}} \end{pmatrix} \begin{pmatrix} 1 & & \\ & 0 & 1 \\ & 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{\sqrt{2}}{2\sqrt{5}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3\sqrt{2}}{2\sqrt{5}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{2\sqrt{2}}{3\sqrt{5}} & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

例 6　化二次型

$$f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 - 6x_2x_3$$

为标准形，并求出其变数变换。

解　因没有平方项，所以不能直接配方，必须想法变出平方项。为此，令变数变换

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \tag{16}$$

代入原式，得

$$f(x_1, x_2, x_3) = g(y_1, y_2, y_3) = 2y_1^2 - 2y_2^2 - 4y_1y_3 + 8y_2y_3. \tag{17}$$

接着按例 5 的办法配方，即

$$\begin{aligned} g(y_1, y_2, y_3) &= 2(y_1^2 - 2y_1 \cdot y_3 + y_3^2) - 2y_2^2 + 8y_2y_3 - 2y_3^2 \\ &= 2(y_1 - y_3)^2 - 2y_2^2 + 8y_2y_3 - 2y_3^2. \end{aligned}$$

令变数变换

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \quad \text{或} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}.$$

代入上式，得

$$\begin{aligned} h(z_1, z_2, z_3) &= 2z_1^2 - 2z_2^2 + 8z_2z_3 - 2z_3^2 \\ &= 2z_1^2 - 2(z_2^2 - 4z_2z_3 + z_3^2) \\ &= 2z_1^2 - 2(z_2 - 2z_3)^2 + 6z_3^2. \end{aligned}$$

令变数变换

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}, \qquad \text{或} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

代入上式，得

$$k(u_1, u_2, u_3) = 2u_1^2 - 2u_2^2 + 6u_3^2.$$

再令

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & & \\ & \frac{1}{\sqrt{2}} & \\ & & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

代入上式，得

$$l(v_1, v_2, v_3) = v_1^2 - v_2^2 + v_3^2.$$

最后令

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

代入上式，得

$$m(t_1, t_2, t_3) = t_1^2 + t_2^2 - t_3^2. \tag{18}$$

故 $f(x_1, x_2, x_3)$ 的标准形为（18），其由 $f(x_1, x_2, x_3)$ 到标准形（18）的变数变换为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & & \\ & \frac{1}{\sqrt{2}} & \\ & & \frac{1}{\sqrt{6}} \end{pmatrix}$$

$$\times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{6}} & 0 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}.$$

从以上两个例子我们看到，将二次型化为标准形，都是按着某一平方项并项配方，然后进行一次变数变换，这次变数变换的实质就是用（7）型阵去演化二次型的表示矩阵。比如，例5一开始就按x_1并项配方，施行变数变换（9），这次变数变换的实质就是利用（9）的系数阵去演化二次型$f(x_1, x_2, x_3)$的表示矩阵，即

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{3}{2} & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 & 3 \\ -2 & 0 & \frac{3}{2} \\ 3 & \frac{3}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & -\frac{3}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -2 & \frac{9}{2} \\ 0 & \frac{9}{2} & -\frac{9}{2} \end{pmatrix}. \qquad (19)$$

显然以（19）的对称阵为表示矩阵的二次型恰好为(10)。

当二次型没有平方项时，施行一次变数变换使得有平方项，而这次变数变换的实质就是用（8）型阵去演化二次型的表示矩阵。比如，例6一开始因没有平方项，而施行变数变换（16）就变为有平方项的二次型（17），这实质就是利用（16）的系数阵去演化二次型$f(x_1, x_2, x_3)$的表示矩阵，即

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -3 \\ 1 & -3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 0 & -2 & 4 \\ -2 & 4 & 0 \end{pmatrix}. \qquad (20)$$

显然以（20）的对称阵为表示矩阵的二次型恰好为（17）．综上所述，按二次型的平方项配方来化简二次型的方法，其实质就是反复用（7）型阵或（8）型阵去化二次型的表示矩阵，使成为对角形，进而成为标准形，这样就得到了用（7）型阵或（8）型阵去化对称阵，使成为对角形，进而成为标准形的方法，因而就称为配方法．

由于可用配方法把对称阵化为对角形，进而化成标准形，所以用配方法化简二次型可不必象例5、例6那样去做，而先写出二次型的表示矩阵．然后用配方法把它化为标准形，于是就得到了这个二次型的标准形了．

例7 用配方法化二次型

$f(x_1,x_2,x_3,x_4)=4x_1x_2+2x_1x_3-6x_1x_4-4x_2x_3+2x_2x_4+2x_3x_4$为标准形．并求出其变数变换．

解 先写出表示矩阵

$$A=\begin{pmatrix} 0 & 2 & 1 & -3 \\ 2 & 0 & -2 & 1 \\ 1 & -2 & 0 & 1 \\ -3 & 1 & 1 & 0 \end{pmatrix}.$$

再用（7）型阵或（8）型阵去化A，由于这个对称阵就是例4那个对称阵，所以用配方法化为标准形：

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}.$$

故二次型$f(x_1,x_2,x_3,x_4)$的标准形为：

$$g(y_1,y_2,y_3,y_4)=y_1^2+y_2^2-y_3^2-y_4^2,$$

其变数变换为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = P\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}, \text{ 即} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{2}{\sqrt{5}} & \frac{1}{2} & \frac{3}{2\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{\sqrt{5}} & -\frac{1}{2} & \frac{1}{2\sqrt{2}} \\ 0 & \frac{2}{\sqrt{5}} & 0 & \frac{2}{\sqrt{2}} \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}.$$

练 习 二

1．用两种方法将下列对称阵化为对角形，并求出相合演化矩阵：

1) $\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$，　2) $\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \end{pmatrix}$.

2．用两种方法将下列对称阵化为标准形，并求出相合演化阵：

1) $\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$，　2) $\begin{pmatrix} 2 & 1 & -1 & -1 \\ 1 & 0 & -2 & -2 \\ -1 & -2 & 1 & 1 \\ -1 & -2 & 1 & 1 \end{pmatrix}$.

3．将下列二次型化为平方和，并求出可逆变数变换：

1) $f(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + 2x_2^2 + 4x_2x_3 + 4x_3^2$，

2) $f(x_1, x_2, x_3, x_4) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$.

4．将下列二次型化为标准形，并求出可逆变数变换：

1) $f(x_1, x_2, x_3) = 2x_1x_2 + 2x_1x_3 + 6x_2x_3$，

2) $f(x_1, x_2, x_3, x_4) = x_1^2 + 2x_1x_2 + x_2^2 + 2x_2x_3 + x_3^2 + 2x_3x_4 + x_4^2$.

5. 证明：秩为 r 的对称阵可以表为 r 个秩为 1 的对称阵之和。

§3 惯性定律

我们已经知道，秩为 r 的对称阵 A 必相合于下列特殊的对角形阵：

$$\begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & -1 & & & & \\ & & & & \ddots & & & \\ & & & & & -1 & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix} = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix} \qquad (1)$$

（其中 1 有 p 个，-1 有 q 个）

其中 $p+q=r$。

本节主要证明对角形阵（1）中正 1 个数 p 与负 1 个数 q 的分配是由对称阵 A 所唯一确定的（不仅其总和 $p+q=r$ 是由 A 所唯一确定）。当中正 1 的个数 p 叫做 A 的正惯标，负 1 的个数 q 叫做 A 的负惯标，它们的差 $s=p-q$ 叫做 A 的符号差。这样本节要证的基本结果是：秩与符号差是由 A 所唯一确定，进而秩与符号差是对称阵在相合分类之下的完备不变量。

我们先证明一个命题，为证明基本结果做准备，同时它本身也是重要的。

命题（*Witt* 定理）设 A 是 n 阶可逆对称阵，B_1, B_2 是两个 m 阶可逆对称阵。如果两个 $n+m$ 阶可逆对称阵

$$\begin{pmatrix} A & 0 \\ 0 & B_1 \end{pmatrix},\ \begin{pmatrix} A & 0 \\ 0 & B_2 \end{pmatrix}.$$

相合，则 B_1 与 B_2 也相合。

证明　依上节定理 1，有可逆阵 Q，使

$$A = Q'\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_2 \end{pmatrix}Q = Q'\mathrm{diag}(\lambda_1, \cdots, \lambda_n)Q.$$

$\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$表示对角线元素为$\lambda_1,\cdots,\lambda_n$的对角阵，于是

$$\begin{pmatrix} A & 0 \\ 0 & B_1 \end{pmatrix} = \begin{pmatrix} Q'\mathrm{diag}(\lambda_1,\cdots\lambda_n)Q & 0 \\ 0 & B_1 \end{pmatrix}$$

$$= \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix}' \begin{pmatrix} \mathrm{diag}(\lambda_1,\cdots\lambda_n) & 0 \\ 0 & B_1 \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix},$$

$$\begin{pmatrix} A & 0 \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} Q'\mathrm{diag}(\lambda_1,\cdots,\lambda_n)Q & 0 \\ 0 & B_2 \end{pmatrix}$$

$$= \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix}' \begin{pmatrix} \mathrm{diag}(\lambda_1,\cdots,\lambda_n) & 0 \\ 0 & B_2 \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix}.$$

因为

$$\begin{pmatrix} A & 0 \\ 0 & B_1 \end{pmatrix} 与 \begin{pmatrix} A & 0 \\ 0 & B_2 \end{pmatrix}$$

相合，所以对称阵

$$\begin{pmatrix} \mathrm{diag}(\lambda_1,\cdots,\lambda_n) & 0 \\ 0 & B_1 \end{pmatrix} 与 \begin{pmatrix} \mathrm{diag}(\lambda_1,\cdots,\lambda_n) & 0 \\ 0 & B_2 \end{pmatrix}$$

相合. 记

$$B_1^* = \begin{pmatrix} \mathrm{diag}(\lambda_2,\cdots,\lambda_n) & 0 \\ 0 & B_1 \end{pmatrix}, B_2^* = \begin{pmatrix} \mathrm{diag}(\lambda_2,\cdots,\lambda_n) & 0 \\ 0 & B_2 \end{pmatrix}.$$

则 B_1^* 与 B_2^*仍为可逆对称阵，且

$$\begin{pmatrix} \lambda_1 & \\ & B_1^* \end{pmatrix} 与 \begin{pmatrix} \lambda_1 & \\ & B_2^* \end{pmatrix}$$

相合. 所以存在 $n+m$ 阶可逆阵R，把R按第一行、第一列分块，记为

$$R = \begin{pmatrix} a & \alpha \\ \beta' & R_1 \end{pmatrix}$$

其中 α,β 是 $1\times(n+m-1)$矩阵，R_1 是 $n+m-1$ 阶方阵. 于是，有

$$\begin{pmatrix}\lambda_1 & \\ & B_2^*\end{pmatrix} = R\begin{pmatrix}\lambda_1 & \\ & B_1^*\end{pmatrix}R' = \begin{pmatrix}a & \alpha \\ \beta' & R_1\end{pmatrix}\begin{pmatrix}\lambda_1 & \\ & B_1^*\end{pmatrix}\begin{pmatrix}a & \beta \\ \alpha' & R'_1\end{pmatrix}$$

$$= \begin{pmatrix}\lambda_1 a^2 + \alpha B_1^*\alpha' & \lambda_1 a\beta + \alpha B_1^* R'_1 \\ \lambda_1 a\beta' + R_1 B_1^*\alpha' & \lambda_1\beta'\beta + R_1 B_1^* R'_1\end{pmatrix}.$$

所以得到关系

$$\begin{cases}\lambda_1(1 - a^2) = \alpha B_1^*\alpha' \\ (\lambda_1 a)\beta + \alpha B_1^* R_1' = 0 \\ \lambda_1\beta'\beta + R_1 B_1^* R_1' = B_2^*.\end{cases}$$

如果 $a=1$，可取 $-R$ 同样讨论之，这时 $-R$ 的 $(1,1)$ 位置元素是 -1，所以不妨假定 $a\neq1$，这样有

$$[R_1 + (1-a)^{-1}\beta'\alpha]B_1^*[R_1 + (1-a)^{-1}\beta'\alpha]'$$

$$= R_1B_1^*R_1' + (1-a)^{-1}(\beta'\alpha B_1^*R_1' + R_1B_1^*\alpha'\beta) + (1-a)^{-2}\beta'\alpha B_1^*\alpha'\beta$$

$$= B_2^* - \lambda_1\beta'\beta + (1-a)^{-1}(-\lambda_1 a\beta'\beta - \lambda_1 a\beta'\beta) + (1-a)^{-2}\lambda_1(1-a^2)\beta'\beta$$

$$= B_2^* + \left(\frac{1+a}{1-a} - \frac{2a}{1-a} - 1\right)\lambda_1\beta'\beta = B_2^*.$$

于是取 $K = R_1 + (1-a)^{-1}\beta'\alpha$，则 $B_2^* = KB_1^*K'$。双方取行列式便知 K 是可逆阵，故 B_1^* 与 B_2^* 相合。同理可证

$$\begin{pmatrix}\mathrm{diag}(\lambda_3,\cdots,\lambda_n) & 0 \\ 0 & B_1\end{pmatrix} 与 \begin{pmatrix}\mathrm{diag}(\lambda_3,\cdots,\lambda_n) & 0 \\ 0 & B_2\end{pmatrix}$$

相合。这样继续讨论下去，最后可证 B_1 与 B_2 相合。证完。

定理（惯性定律） 如果两个 n 阶对称方阵

$$\begin{pmatrix}I_p & & \\ & -I_q & \\ & & 0\end{pmatrix} 与 \begin{pmatrix}I_{p_1} & & \\ & -I_{q_1} & \\ & & 0\end{pmatrix}$$

相合，则 $p+q=p_1+q_1$，且 $p=p_1$，$q=q_1$（从而 $s=p-q=p_1-q_1=s_1$）。

证明 $p+q=p_1+q_1$ 是明显的。剩下证 $p=p_1$，$q=q_1$。我们取可逆阵 Q，将 Q 按前 $r(=p+q)$ 行，前 r 列分块，即

$$Q=\begin{pmatrix}Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{pmatrix}$$

并且使

$$\begin{pmatrix}I_p & & \\ & -I_q & \\ & & 0\end{pmatrix}=Q\begin{pmatrix}I_{p_1} & & \\ & -I_{q_1} & \\ & & 0\end{pmatrix}Q'$$

$$=\begin{pmatrix}Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{pmatrix}\begin{pmatrix}\begin{pmatrix}I_{p_1} & \\ & -I_{q_1}\end{pmatrix} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}Q'_{11} & Q'_{21}\\ Q'_{12} & Q'_{22}\end{pmatrix}.$$

于是

$$\begin{pmatrix}I_p & \\ & -I_q\end{pmatrix}=Q_{11}\begin{pmatrix}I_{p_1} & \\ & -I_{q_1}\end{pmatrix}Q'_{11}.$$

双方取行列式，可知 $\det Q_{11}\neq 0$。故

$$\begin{pmatrix}I_p & \\ & -I_q\end{pmatrix}\text{与}\begin{pmatrix}I_{p_1} & \\ & -I_{q_1}\end{pmatrix}$$

相合。

假设$p\neq p_1$，不妨设 $p>p_1$，由 Witt 定理可知

$$\begin{pmatrix}I_{p-p_1} & \\ & -I_q\end{pmatrix}\text{与}-I_{q_1}$$

相合。所以有 $r-p_1$ 阶可逆阵R，使

$$\begin{pmatrix}I_{p-p_1} & \\ & -I_q\end{pmatrix}=R(-I_{q_1})R'=-RR'$$

即
$$\begin{pmatrix}-I_{p-p_1} & \\ & I_q\end{pmatrix}=RR'.$$

由于R是可逆的，所以它的第一行元素 $b_{11},\cdots,b_{1\,r-p_1}$ 是不全为 0 的实数。但由上式可得

$$b_{11}^2+\cdots+b_{1r-p_1}^2=-1$$

这是不可能的，故p不能大于 p_1。同理可证 p 不能小于 p_1。因此

必有 $p=p_1$。证完。

惯性定律表明对称阵在相合分类之下的标准形（1）是唯一的。把这一结论应用于二次型便有：任意实二次型

$$f(x_1,\cdots,x_n)=(x_1\cdots x_n)A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

都可用适当的可逆变数变换

$$\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=Q\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix},\ detQ\neq 0$$

化简为标准形：

$$f(x_1,\cdots,x_n)=g(y_1,\cdots,y_n)=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2 \quad (2)$$

当中的 p,q，$p+q$ 均由二次型 $f(x_1,\cdots,x_n)$ 唯一确定，分别叫做**二次型的正惯标，负惯标和秩**，而 $s=p-q$ 叫做符号差。这里明显看出二次型的秩就是它的表示矩阵的秩。

下面简要的说一下关于复数域上的对称阵的相合问题。与实数域上对称阵的情形完全一样，可以证明：**任意复对称阵 A 都能用成套的初等变换化成对角形阵**

$$\begin{pmatrix}\lambda_1 &&&&&\\ &\lambda_2&&&&\\ &&\ddots&&&\\ &&&\lambda_r&&\\ &&&&0&\\ &&&&&\ddots\\ &&&&&&0\end{pmatrix}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 均是非零复数。

因为非零复数总可以开平方，于是在化得上述对角形的基础上，再用一些成套的初等变换可化为

$$\begin{pmatrix}1&&&&\\ &\ddots&&&\\ &&1&&\\ &&&0&\\ &&&&\ddots\\ &&&&&0\end{pmatrix}\ (r\text{个}1).$$

这说明复对称阵化成的标准形是由该阵的秩数所唯一确定. 即秩数是复对称阵在相合分类之下的完备不变量.

练　习　三

1. 如果把 n 阶对称阵按相合关系分类, 即两个 n 阶对称阵属于同一类必要而且只要它们是相合的, 试问一共能分几类?

2. 证明: 对称阵的秩 r 与符号差 $s=p-q$ 是同奇偶的, 并且 $|s|\leqslant r$, 其中 p 为正惯标, q 为负惯标.

3. 设 A 为实对称阵, 且 $|A|<0$. 证明 A 的负惯标必是一个正数.

4. 设实二次型 $f(x_1,\cdots,x_n)$ 的正负惯标分别为 p,q; $a_1,\cdots,a_p$ 是任 p 个正数, $b_1,\cdots,b_q$ 是任 q 个负数. 那么这个二次型可以化成

$$g(y_1,\cdots,y_n)=a_1y_1^2+\cdots+a_py_p^2+b_1y_{p+1}^2+\cdots+b_qy_{p+q}^2.$$

§4　正定矩阵

本节讨论一类特殊的实对称阵——正定矩阵. 它是实二次型(做为二次齐次函数)取值规律的一种反映.

设 A 为任一 n 阶实对称阵, 那么 A 必相合于它的标准形:

$$\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}$$

其中 $p+q=r$, r 为 A 的秩数.

定义 1　如果 A 的正惯标 $p=A$ 的秩 $r=A$ 的阶数 n, 则称 A 为正定的;

如果 A 的负惯标 $q=A$ 的秩 $r=A$ 的阶数 n, 则称 A 为负定的;

如果 A 的正惯标 $p=A$ 的秩数 $r\leqslant A$ 的阶数 n, 则称 A 为半

正定的；

如果A的负惯标 $q=A$的秩数 $r\leqslant A$的阶数 n，则称A为半负定的。

例如，阵

$$A=\begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix},\quad B=\begin{pmatrix}-1&&\\&-1&\\&&-1\end{pmatrix},\quad C=\begin{pmatrix}1&&\\&1&\\&&0\end{pmatrix},$$

$$D=\begin{pmatrix}-1&&\\&0&\\&&0\end{pmatrix},$$

中A为正定的，B为负定的，C为半正定的，D为半负定的。

明显的有：A为正（负）定的必要而且只要A相合于$I_n(-I_n)$。

命题1　A为（半）正定的必要而且只要$-A$为（半）负定的。

结论的正确性是明显的。依此命题，我们只须讨论（半）正定矩阵即可。

命题2　A为正定的充分必要条件是：对任意非零 $n\times 1$ 矩阵

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

恒有

$$X'AX>0;$$

A为半正定的充分必要条件是：对任意非零 $n\times 1$矩阵

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

恒有

$$X'AX\geqslant 0.$$

证明　$X'AX$ 为一个二次型，令

$$f(x_1,\cdots,x_n)=X'AX.$$

再令Q化A成标准形：

$$Q'AQ=\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}. \tag{1}$$

于是，设

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}=Q\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \tag{2}$$

则有

$$f(x_1,\cdots,x_n)=X'AX=(y_1\cdots y_n)Q'AQ\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2. \tag{3}$$

这时，如果A为正定的，那么A的正惯标 $p=A$的秩 $r=A$的阶数 n。就是说（3）中的 $p=n$，$q=0$。于是当 $X\neq 0$ 时，由（2），必有

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\neq 0$$

故由（3）知，$X'AX>0$。反之，由非零X，恒有$X'AX>0$，则必定（3）中的 $p=n$。这因为若 $p<n$，取

$$\begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{4}$$

代入（3）后的值就小于或等于0了。将（4）代入（2）就得到非零X，有 $X'AX\leqslant 0$，这与题设矛盾；

类似地，如果 A 为半正定的，那么A的正惯标 $p=A$ 的秩 $r\leqslant A$的阶数 n。就是说（3）中的 $p=r\leqslant n$，$q=0$。于是当 $X\neq 0$，由（2），必有

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \neq 0$$

故由（3）知， $X'AX \geqslant 0$．反之，由非零X，恒有 $X'AX \geqslant 0$，则必定（3）中的 $p = r \leqslant n$，$q = 0$．不然若 $q > 0$ 取

$$\begin{pmatrix} y_1 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

代入（3）就产生了矛盾．证完．

命题2说明了对称阵A为（半）正定的意义，即对称阵的（半）正定性反映了由它做表示矩阵的二次型$X'AX$的取值规律．

命题3　A为正定的充分必要条件是：有可逆阵 P，使 $A = P'P$．

证明　充分性．由 $A = P'P = P'I_nP$，说明A相合于单位阵 I_n，故A为正定的．

必要性，设A为正定的，则A与单位阵 I_n 相合，所以存在可逆阵 P，使 $A = P'I_nP = P'P$，证完．

推论　正定矩阵A的行列式 $\det A > 0$．

命题4　A为半正定的充分必要条件是：有方阵 Q，使 $A = Q'Q$．

证明　必要性，由A为半正定，有可逆阵P，使

$$A = P' \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} P \quad (r \text{个} 1)$$

$$=P'\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}P$$

（其中 1 共 r 个）

令

$$Q=\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}P,$$

则有

$$A=Q'Q.$$

充分性。由 $A=Q'Q$，故对任意非零的$n\times 1$矩阵

$$X=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

都有 $X'AX=X'Q'QX=(QX)'(QX)$，令

$$QX=\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

故

$$X'AX=(y_1\cdots y_n)\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}=y_1^2+\cdots+y_n^2\geqslant 0.$$

所以A为半正定。证完。

以上讨论了（半）正定矩阵的性质，下面给出（半）正定矩阵的重要判定条件。为此，先给出

定义2 设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

是一n阶方阵，行标和列标相同的子块

$$\begin{pmatrix} a_{i_1i_1} & a_{i_1i_2} & \cdots & a_{i_1i_k} \\ a_{i_2i_1} & a_{i_2i_2} & \cdots & a_{i_2i_k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_ki_1} & a_{i_ki_2} & \cdots & a_{i_ki_k} \end{pmatrix} (1\leqslant i_1<i_2<\cdots<i_k\leqslant n)$$

叫做A的主子块；而子式

$$\begin{vmatrix} a_{i_1i_1} & a_{i_1i_2} & \cdots & a_{i_1i_k} \\ a_{i_2i_1} & a_{i_2i_2} & \cdots & a_{i_2i_k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_ki_1} & a_{i_ki_2} & \cdots & a_{i_ki_k} \end{vmatrix}$$

叫做A的主子式，记作 $D(i_1, i_2, \cdots, i_k)$。而子式

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix} (k=1, 2, \cdots, n)$$

叫做A的顺序主子式。

例如，设$A=\begin{pmatrix} 1 & -1 & 2 & 3 \\ -1 & 0 & -1 & 1 \\ 2 & -1 & 2 & 0 \\ 3 & 1 & 0 & -1 \end{pmatrix}$。

那么

$$1, 2, \begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix}, \begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix}, \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix}, \begin{vmatrix} 1 & -1 & 3 \\ -1 & 0 & 1 \\ 3 & 1 & -1 \end{vmatrix},$$

$$\begin{vmatrix} 0 & -1 & 1 \\ -1 & 2 & 0 \\ 1 & 0 & -1 \end{vmatrix}$$

都是A的主子式；而A的顺序主子式则只有4个，即

$$1, \quad \begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & -1 & 2 \\ -1 & 0 & -1 \\ 2 & -1 & 2 \end{vmatrix}, \quad \begin{vmatrix} 1 & -1 & 2 & 3 \\ -1 & 0 & -1 & 1 \\ 2 & -1 & 2 & 0 \\ 3 & 1 & 0 & -1 \end{vmatrix}.$$

定理1 n **阶对称阵**

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

是正定的充分必要条件为它的 n**个顺序主子式**

$$|a_{11}| > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \cdots, \begin{vmatrix} a_{11} & \cdots & a_{1n-1} \\ \cdots & \cdots & \cdots \\ a_{n-11} & \cdots & a_{n-1n-1} \end{vmatrix} > 0, \quad |A| > 0.$$

证明 必要性考察阵A的任一k阶顺序主子块

$$A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \cdots & \cdots & \cdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

它本身做为k阶对称阵必是正定的，否则有非零的$k\times 1$矩阵

$$\alpha = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}$$

使

$$\alpha' A_k \alpha \leqslant 0.$$

于是取非零的$n\times 1$矩阵

$$X=\begin{pmatrix} c_1 \\ \vdots \\ c_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

便有

$$X'AX=(\alpha'\ 0)\begin{pmatrix} A_k & * \\ * & * \end{pmatrix}\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$$

$$=(\alpha' A_k\ *)\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$$

$$=\alpha' A\alpha$$

$$\leqslant 0,$$

即 $$X'AX\leqslant 0.$$

这与A为正定阵矛盾，由命题3的推论知，有

$$\det A_k>0,\quad 1\leqslant k\leqslant n.$$

充分性。对n用数学归纳法。

当 $n=1$时， $A=(a_{11})$，即 $|a_{11}|=a_{11}>0$，这时显然A是正定的。

假设对 $n-1$ 阶对称阵定理1成立。往证对n阶对称阵A定理1成立。

事实上，将A按前 $n-1$行、$n-1$ 列分块，即

$$A=\begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix},\quad \alpha=\begin{pmatrix} a_{1n} \\ \vdots \\ a_{n-1n} \end{pmatrix},\quad A_{n-1}=\begin{pmatrix} a_{11} & \cdots & a_{1n-1} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{n-11} & \cdots & a_{n-1n-1} \end{pmatrix}.$$

于是 A_{n-1} 是 $n-1$ 阶对称阵，它的 $n-1$ 个顺序主子式

$$|a_{11}|>0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\quad \cdots,\quad \begin{vmatrix} a_{11} & \cdots & a_{1n-1} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{n-11} & \cdots & a_{n-1\,n-1} \end{vmatrix}>0.$$

根据归纳法假设，A_{n-1} 是正定的，因而 A_{n-1} 相合于单位阵I_{n-1}。即有 $n-1$ 阶可逆阵Q，使

$$Q'A_{n-1}Q = I_{n-1}.$$

令

$$Q_1 = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix},$$

则 Q_1 是 n 阶可逆阵，且有

$$Q_1'AQ_1 = \begin{pmatrix} Q' & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} A_{n-1} & \alpha \\ \alpha' & a_{nn} \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} Q'A_{n-1}Q & Q'\alpha \\ \alpha'Q & a_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} I_{n-1} & Q'\alpha \\ \alpha'Q & a_{nn} \end{pmatrix} = B.$$

这个等式说明，用可逆阵 Q 将 A 化为 B 了。在此基础上，用可逆阵

$$Q_2 = \begin{pmatrix} I_{n-1} & -Q'\alpha \\ 0 & 1 \end{pmatrix}$$

继续化简 A，也就是化简 B，得

$$Q_2'BQ_2 = \begin{pmatrix} I_{n-1} & 0 \\ -\alpha'Q & 1 \end{pmatrix}\begin{pmatrix} I_{n-1} & Q'\alpha \\ \alpha'Q & a_{nn} \end{pmatrix}\begin{pmatrix} I_{n-1} & -Q'\alpha \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} I_{n-1} & Q'\alpha \\ 0 & -\alpha'QQ'\alpha + a_{nn} \end{pmatrix}\begin{pmatrix} I_{n-1} & -Q'\alpha \\ 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} I_{n-1} & 0 \\ 0 & a_{nn} - \alpha'QQ'\alpha \end{pmatrix}.$$

令 $\qquad P_1 = Q_1Q_2,\ \ a_{nn} - \alpha'QQ'\alpha = a.$

则有

$$P_1'AP_1 = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \\ & & & a \end{pmatrix}.$$

由 A 的行列式，$\det A > 0$，推知 $a > 0$，于是再用一套倍法变换 $D_n\left(\frac{1}{\sqrt{a}}\right)$，就得

$$D_n\left(\frac{1}{\sqrt{a}}\right)P_1'AP_1D_n\left(\frac{1}{\sqrt{a}}\right)=I_n.$$

令 $$P=P_1D_n\left(\frac{1}{\sqrt{a}}\right),$$

则有 $$P'AP=I_n.$$

故A为正定的。证完。

推论 n 阶对称阵A为正定的充分必要条件是：A 的一切主子式都是正的：

$$D(i_1,\cdots,i_k)=\begin{vmatrix}a_{i_1i_1} & \cdots & a_{i_1i_k}\\ \cdots & \cdots & \cdots\\ a_{i_ki_1} & \cdots & a_{i_ki_k}\end{vmatrix}>0.$$

事实上，A的主子式$D(i_1,\cdots,i_k)$可经成套的换法变换变为左上角，成为与A相合的新阵 B 的顺序主子式，故有$D(i_1,\cdots,i_k)>0$.

定理 2 n 阶对称阵 $A=(a_{ij})$ 是半正定的充分必要条件是：A 的一切主子式

$$D(i_1,\cdots,i_k)=\begin{vmatrix}a_{i_1i_1} & \cdots & a_{i_1i_k}\\ \cdots & \cdots & \cdots\\ a_{i_ki_1} & \cdots & a_{i_ki_k}\end{vmatrix}\geqslant 0$$

$$1\leqslant i_1<i_2<\cdots<i_k\leqslant n.$$

证明 必要性．考虑A的任一k阶主子块

$$A_k=\begin{pmatrix}a_{i_1i_1} & \cdots & a_{i_1i_k}\\ \cdots & \cdots & \cdots\\ a_{i_ki_1} & \cdots & a_{i_ki_k}\end{pmatrix}$$

它本身做为k阶对称阵必是半正定的，否则有非零的 $k\times 1$ 矩阵

$$\alpha=\begin{pmatrix}c_1\\ \vdots\\ c_k\end{pmatrix}$$

使

$$\alpha' A_k \alpha < 0 .$$

于是取非零的 $n\times 1$ 矩阵

$$X=\begin{pmatrix} c_1 \\ \vdots \\ c_k \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

便有

$$\begin{aligned} X'BX &= (\alpha'\ 0)\begin{pmatrix} A_k & * \\ * & * \end{pmatrix}\begin{pmatrix} \alpha \\ 0 \end{pmatrix} \\ &= (\alpha' A_k\ \ *)\begin{pmatrix} \alpha \\ 0 \end{pmatrix} \\ &= \alpha' A_k \alpha \\ &< 0 . \end{aligned}$$

即$X'BX<0$。故 B 不是半正定的，那么A也不是半正定的，这是一个矛盾，其中 B 是A经过适当的成套换法变换，使 A_k 变为顺序主子式的新阵 B。由命题 4 知，半正定矩阵的行列式是非负的。故

$$\det A_k \geqslant 0,\ 1\leqslant k\leqslant n.$$

充分性。任取A的 k阶顺序主子块

$$A_k=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix},\quad (k=1,2,\cdots,n).$$

于是k 阶行列式

$$|\lambda I_k + A_k| = \lambda^k + b_1\lambda^{k-1} + \cdots + b_{k-1}\lambda + b_k.$$

经过计算可知，右端系数 b_i 等于 A_k 的一切 i 阶主子式之和。因此对任意取定的 $\lambda_0>0$，恒有

$$|\lambda_0 I_k + A_k| > 0,\ k=1,2,\cdots,n.$$

据上面的定理 1，n 阶对称阵 $\lambda_0 I_n + A$ 是正定的。有了这一段说明之后，我们用反证法来证明充分性。

假定A不是半正定的，必有 $n\times 1$ 矩阵 $X\neq 0$，使

$$X'AX=-c,\quad c>0.$$

于是取 $\lambda_0=\frac{c}{X'X}$，则 $\lambda_0>0$，且$\lambda_0 X'X=c$。

我们看 n阶对称阵 $B=\lambda_0 I_n+A$，有

$$X'BX=X'(\lambda_0 I_n+A)X=\lambda_0 X'X+X'AX=c-c=0,$$

这与 $B=\lambda_0 I_n+A$ 是正定的矛盾，故A为半正定的。证完。

最后，与正定矩阵相应的，用二次型语言给出

定义 3 设实二次型

$$f(x_1,\cdots,x_n)=(x_1\cdots x_n)A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}.$$

若表示矩阵A为正定的，则称二次型 $f(x_1,\cdots,x_n)$ 为正定的；

若表示矩阵A为负定的，则称二次型 $f(x_1,\cdots,x_n)$ 为负定的；

若表示矩阵A为半正定的，则称二次型 $f(x_1,\cdots,x_n)$ 为半正定的；

若表示矩阵A为半负定的，则称二次型 $f(x_1,\cdots,x_n)$ 为半负定的。

关于（半）正定矩阵所得到的结果，相应的二次型也都有此结果，这里不再重述了。

例 1 判别对称阵

$$A=\begin{pmatrix}3&2&0\\2&4&-2\\0&-2&5\end{pmatrix}$$

是否正定。

解 A的顺序主子式

$$3>0,\quad \begin{vmatrix}3&2\\2&4\end{vmatrix}=8>0,\quad \begin{vmatrix}3&2&0\\2&4&-2\\0&-2&5\end{vmatrix}=28>0.$$

据定理1，A为正定的。

例2 判别对称阵

$$A=\begin{pmatrix}1&1&2\\1&1&2\\2&2&1\end{pmatrix}$$

是否半正定。

解 因A的一个二阶主子式

$$\begin{vmatrix}1&2\\2&1\end{vmatrix}=-3<0.$$

据定理2知，A不是半正定的。

例3 判别二次型

$$f(x_1,x_2,x_3)=3x_1^2+4x_2^2+5x_3^2+4x_1x_2-4x_2x_3$$

是否正定。

解 二次型$f(x_1,x_2,x_3)$的表示矩阵

$$A=\begin{pmatrix}3&2&0\\2&4&-2\\0&-2&5\end{pmatrix}.$$

由上例1知，A为正定的，故二次型$f(x_1,x_2,x_3)$也是正定的。

练 习 四

1. 判别下列对称阵是否正定：

1) $\begin{pmatrix}1&1&1\\1&2&3\\1&3&6\end{pmatrix}$， 2) $\begin{pmatrix}1&2&-2\\2&4&-1\\-2&-1&0\end{pmatrix}$， 3) $\begin{pmatrix}-5&2&2\\2&-6&0\\2&0&-4\end{pmatrix}$，

4) $\begin{pmatrix}1&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&1&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&1&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&\frac{1}{2}&1\end{pmatrix}$.

2. 判别下列二次型是否正定：

1) $x_1^2+2x_1x_2+2x_2^2+4x_2x_3+5x_3^2$；

2) $5x_1^2+6x_2^2+4x_3^2-4x_1x_2-4x_2x_3$；

3) $x_1^2+x_2^2+4x_3^2+7x_4^2+6x_1x_3+4x_1x_4-4x_2x_3+2x_2x_4+4x_3x_4$。

3. 假如对任意 $n\times1$ 矩阵

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix},\quad x_1\neq0,\ \cdots,\ x_n\neq0$$

都有 $X'AX>0$，试问A是否正定？

4. 若A为对称阵，M为可逆阵，则 $M'AM$ 与A的正定、负定、半正定、半负定性是一致的。

5. 若A是可逆阵，则 $A'A$ 为正定的。

6. 若 A、B 都是 n 阶正定阵，则 $A+B$，也是正定阵。

7. 设A为对称阵，且 $|A|<0$。证明必存在实 $n\times1$ 矩阵 $X\neq0$，使 $X'AX<0$

习 题 九

1. 设 A、B 皆为对称阵，证明AB为对称阵的充分必要条件是：$AB=BA$。

2. 设 $f(x)$ 为实系数多项式，A为对称阵．证明 $f(A)$ 为对称阵。

3. 若A为正定的，则 A^* 也为正定的。

4. 若$A=(a_{ij})$为正定的，则 $a_{ii}>0$，$(i=1,2,\cdots,n)$。

5. 设 $A=(a_{ij})$ 为正定的，$b_1,\cdots,b_n$ 为任 n 个非零实数．证明 $B=(a_{ij}b_ib_j)$ 也是正定的。

6. 设A为对称阵．证明当实数 t 充分大后，tI_n+A 为正定的。

7. 证明：$X'AX=n\Sigma x_i^2-(\Sigma x_i)^2$是半正定的，其中

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}.$$

8　问当 t 取何值时，二次型

$$f(x_1,x_2,x_3)=x_1^2+4x_2^2+x_3^2+2tx_1x_2+10x_1x_3+6x_2x_3$$

为正定的。

9　设 A 为对称阵。证明：若对任一 $n\times1$ 矩阵 $X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$，均有 $X'AX=0$，则 $A=0$。

10　设 A 为对称阵。证明：只有 $n\times1$ 矩阵 $X=0$ 时，才有 $X'AX=0$，那么 A 必为满秩的。

11　设 A 为对称阵。若有 $n\times1$ 矩阵 X_1,X_2，使

$$X_1'AX_1>0,\quad X_2'AX_2<0.$$

则必存在 $n\times1$ 矩阵 $X_0\neq0$，使 $X_0'AX_0=0$。

第十章 方阵在相似之下的标准形

§1 方阵的相似及相似分类

以上讨论了两种类型的矩阵分类和标准形问题：相抵与相合．它们的矩阵表述形式是

相抵：两个$m\times n$矩阵A，B叫做相抵的充分必要条件是存在m阶可逆阵P及n阶可逆阵Q，使 $B=PAQ$．

相合：两个 n 阶方阵A，B叫做相合的充分必要条件是存在 n 阶可逆阵P，使$B=P'AP$．

相抵的标准形理论全面解决了线性方程组的求解问题和解的结构问题；相合的标准形理论解决了二次型的化简问题（即实对称阵的化简问题）．

在线性变换的矩阵表示和常系数线性微分方程组的求解问题中要遇到这样的矩阵问题：对n阶方阵A是否存在 n 阶可逆阵 P，使$P^{-1}AP$ 具有尽可能简单的形式．容易指出，这正是我们所说的另一种矩阵的分类和标准形问题．为了说明这一事实，首先给出

定义 设A,B为任二n阶方阵．如果存在 n 阶可逆矩阵P，使$B=P^{-1}AP$，则称A与B相似，记作 $A\sim B$，其中的矩阵 P 叫做使 A 与B相似的演化矩阵．**

容易证明，n 阶方阵的相似关系满足能够做成分类的三个条件，即相似关系是等价关系：

1）**反身性** 对任一n 阶方阵A，都有 $A\sim A$；

2）**对称性** 若 $A\sim B$，则 $B\sim A$；

3）**传递性** 若$A \sim B$，$B \sim C$，则 $A \sim C$。

事实上，对任一 n 阶方阵A，取n阶单位阵I_n为演化矩阵，便有$A = I_n^{-1}AI_n$，所以 $A \sim A$；若 $A \sim B$，则有可逆矩阵 P，使 $B = P^{-1}AP$。于是 $A = PBP^{-1} = (P^{-1})^{-1}BP^{-1}$，这表明以 P^{-1} 为演化矩阵使B与A相似：$B \sim A$，即对称性成立；再若 $A \sim B$，$B \sim C$，则有可逆阵P，Q使$B = P^{-1}AP$，$C = Q^{-1}BQ$。于是，$C = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)$，这表明以 PQ 为演化矩阵，使A与 C 相似：$A \sim C$，故传递性成立。

这样，对于确定的 n，全体 n 阶方阵按相似关系做成了分类：n阶方阵A，B分在一类充分必要条件是$A \sim B$。

例如，以2阶方阵为例，按相似关系，全体二阶方阵被分成了类——相似类。在这样分类之下，二阶零阵

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

所在的相似类里显然只含这一个零阵A。类似地，二阶单位阵

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

所在的相似类里也只含这一个二阶单位阵 I_2。可以进一步想一想：还有没有这样的方阵 C，它所在的相似类里只含有一个二阶阵就是C呢？

本章讨论 n 阶方阵在相似分类之下的不变量和标准形问题。

n 阶方阵在相似之下的标准形理论比相抵、相合之下的标准形问题要复杂一些。这一方面由于相似之下的不变量比较复杂也相当隐蔽，从而要想确定相似完备不变量和相似标准形就比较费事；另一方面，由于相似关系与相抵、相合不同，它不能直接使用大家已经熟悉的初等变换方法，来实现相似化简，从而找出相似完备不变量和标准形。但是，我们将会看到，相似完备不变量和相似标准形问题的彻底解决，最终还是少不了以往行之有效的初等变换方法。只不过是经过一段间接迂回的过程罢了。

以下先从一种特殊情况入手，分析n阶方阵与对角形阵相似的条件，并证明实对称阵都相似于对角形阵；接着顺便解决了一类特殊矩阵——正交矩阵的相似标准形问题；最后介绍两种比较普遍的相似标准形的问题和确定相似分类的完备不变量。

练 习 一

1．证明：若 $A\sim B$，则 $A'\sim B'$。

2．证明：若A与B可交换，则 $P^{-1}AP$ 与 $P^{-1}BP$ 也可交换，其中P为可逆阵。

3．证明：若A可逆，则 $AB\sim BA$。

4．设 $A\sim C$，$B\sim D$。证明：

$$\begin{pmatrix} A & \\ & B \end{pmatrix}\sim\begin{pmatrix} C & \\ & D \end{pmatrix}.$$

5．设

$$A=\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix},\quad B=\begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & \ddots & \\ & & & b_n \end{pmatrix}$$

如果 $b_1,b_2,\cdots,b_n$是 $a_1,a_2,\cdots,a_n$的另一种排法，证明：$A\sim B$。

6．证明：n阶方阵A 所在的相似类里只有A本身的充分必要条件是 $A=aI_n$。

7．试问：

1）两个 n阶方阵相抵，它们能否相合？能否相似？

2）两个n阶方阵相似，它们能否相抵？能否相合？

§2　特征矩阵　特征向量

本节讨论方阵A相似于对角形阵的条件。

如果方阵

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

相似于对角形阵

$$B=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

即存在可逆阵P，使

$$P^{-1}AP=\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}. \qquad (1)$$

这里明显地有两个问题必须解决：

1） 什么样的方阵A能相似于对角形阵，即对怎样的方阵A，存在可逆阵P，使（1）成立。如果使（1）成立的P存在的话，如何具体求出来；

2） A所相似的对角形阵B中对角线上的数$\lambda_1, \lambda_2, \cdots, \lambda_n$与$A$有何关系，它们是否是唯一确定的。

为了解决这两个问题，我们来分析（1）式。首先把（1）式改写如下：

$$AP=P\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}. \qquad (2)$$

把（2）式左端的P，右端的对角形阵按列分块，即令

$$P=(\alpha_1, \alpha_2, \cdots, \alpha_n), \quad \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}=(\overline{\lambda_1} \quad \overline{\lambda_2} \cdots \overline{\lambda_n}).$$

于是（2）式可以写成：

$$A(\alpha_1\alpha_2\cdots\alpha_n)=P(\overline{\lambda_1} \quad \overline{\lambda_2} \quad \cdots\overline{\lambda_n}).$$

这样便可分开写出下列n个等式：

$$A\alpha_i = P\overline{\lambda_i},\ i = 1, 2, \cdots, n. \tag{3}$$

因为

$$\overline{\lambda_i} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{pmatrix} (i\text{行})$$

所以由（3）式就有

$$A\alpha_i = \lambda_i\alpha_i,\quad i = 1, 2, \cdots, n. \tag{4}$$

其次，把（4）式左端移于右端，便有

$$(\lambda_i I - A)\alpha_i = 0. \tag{5}$$

形式如（5）的等式，我们是熟知的。这里我们注意到 α_i 是可逆阵P的第i列，所以（5）式表明可逆阵P的第i列α_i是以$\lambda_i I - A$为系数阵的齐次方程组

$$(\lambda_i I - A)X = 0 \tag{6}$$

的解，而由于 α_i 是可逆阵P的一个列，所以是非零解。又齐次方程组（6）有非零解的充分必要条件是系数阵的行列式等于零：

$$|\lambda_i I - A| = 0.$$

为了看得清楚，把上式左端具体写出来就有

$$|\lambda_i I - A| = \begin{vmatrix} \lambda_i - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda_i - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda_i - a_{nn} \end{vmatrix} = 0 \tag{7}$$

这个由A所决定的行列式展开之后经整理是 λ_i 的 n 次多项式

$$\begin{vmatrix} \lambda_i - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda_i - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda_i - a_{nn} \end{vmatrix} = \lambda_i^n + a_1\lambda_i^{n-1} + \cdots + a_{n-1}\lambda_i + a_n.$$

于是（7）式表明：λ_i 是多项式

$$|\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}$$

$$=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n \qquad (8)$$

的一个根。

把以上分析归纳起来，便有——

如果对n阶方阵$A=(a_{ij})$，存在可逆阵 P，使

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

那么

1) $\lambda_1,\lambda_2,\cdots,\lambda_n$ 恰是n次多项式

$$|\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$$

的n 个根；

2) 阵P的第i列 α_i 是齐次线性方程组

$$(\lambda_i I-A)X=0$$

的非零解，且P的 n个列是n个线性无关的解。

这样 n阶方阵A能否相似于对角形阵的问题，就归之于研究由A所决定的多项式（8）和齐次线性方程组（6）。于是，自然地引出以下的特征矩阵，特征向量的讨论。

定义1 设 $A=(a_{ij})$为任一n阶方阵，λ为一个参数。由 A唯一决定的n阶方阵

$$\lambda I-A=\begin{pmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{pmatrix} \qquad (9)$$

叫做A的特征矩阵。特征矩阵的行列式

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

$$= \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n \tag{10}$$

是参数λ的n次多项式，叫做A的特征多项式．A的特征多项式的根叫做A的特征根或特征值．

定义 2　设 λ_i 是A的一个特征根．齐次线性方程组

$$(\lambda_i I - A)X = 0 \tag{11}$$

叫做A的属于 λ_i 的特征方程组．(11)的解空间叫做A的属于 λ_i的特征空间，记作 $S_A(\lambda_i)$，其中的每一个非零解都叫做A的属于λ_i的特征向量①

我们已经见到，式（11）的另一等效写法是

$$AX = \lambda_i X. \tag{12}$$

于是，非零的 $n \times 1$矩阵

$$\alpha = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

是A的属于 λ_i 的特征向量必要而且只要

$$A\alpha = \lambda_i \alpha.$$

按定义直接可得

命题　若A与B相似，则A与B的特征多项式相同，从而特征根也完全一样．

证明　由$A \sim B$，存在可逆阵P，使 $B = P^{-1}AP$．于是

$$|\lambda I - B| = |\lambda I - P^{-1}AP| = |P^{-1}(\lambda I - A)P| = |P^{-1}|\,|\lambda I - A|\,|P| = |\lambda I - A|.$$

---

① 式(11)做为线性方程组，它的解是一组横写的 n 个数，即是一个 $1\times n$ 矩阵。如果把(11)看做是矩阵的式子，它的解就是一个 $n\times 1$矩阵，今后也称其非零解为 A 的特征向量，与前者不予以区别。

说明A与B有相同的特征多项式，当然也有完全一样的特征根，证完。

下面通过具体例子，再进一步熟悉一下特征多项式、特征根、特征向量、特征空间等概念。

例1　设

$$A=\begin{pmatrix}2 & 3 & 2\\ 1 & 4 & 2\\ 1 & -3 & 1\end{pmatrix}.$$

求A的特征根及相应的特征空间、特征向量。

解　首先求出A的特征多项式如下：

$$|\lambda I-A|=\begin{vmatrix}\lambda-2 & -3 & -2\\ -1 & \lambda-4 & -2\\ -1 & 3 & \lambda-1\end{vmatrix}=\lambda^3-6\lambda^2+15\lambda-9.$$

其次求A的特征根，即特征多项式的根：

$$\lambda^3-6\lambda^2+15\lambda-9=(\lambda-1)(\lambda-3)^2$$

所以A的特征根为1，3．其中1是单根，而3是二重根。

再次，分别求出属于A的各特征根的特征空间、特征向量：

1) $S_A(1)$表示齐次线性方程组（即A的属于特征根1的特征方程组）

$$\begin{pmatrix}-1 & -3 & -2\\ -1 & -3 & -2\\ -1 & 3 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

的解空间．经计算知其系数阵的秩为2，所以$S_A(1)$是一维的。这样任取属于1的特征方程组的一个非零解，比如

$$\alpha=(3,1,-3)$$

便构成$S_A(1)$的基低．于是$S_A(1)=L(\alpha)$，从而属于特征根1的特征向量即$S_A(1)$的非零向量可表为

$$k\alpha=k(3,1,-3)=(3k,k,-3k),$$

其中k为任意不等于零的数。

2) $S_A(3)$表示齐次线性方程组（即A的属于特征根3的特征方程组）

$$\begin{pmatrix} 1 & -3 & -2 \\ -1 & -1 & -2 \\ -1 & 3 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的解空间．经计算知其系数阵的秩为2，所以$S_A(3)$也是一维的．这样任取 $S_A(3)$的一个非零向量

$$\beta=(-1,\ -1,1)$$

便构成 $S_A(3)$的基底，于是 $S_A(3)=L(\beta)$。从而属于特征根3的特征向量可表为

$$k\beta=k(-1,\ -1,1)=(-k,\ -k,k),\quad k\neq 0.$$

例2 设

$$B=\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}.$$

求B的特征根及相应的特征向量．

解 首先求出B的特征多项式如下：

$$|\lambda I-B|=\begin{vmatrix} \lambda-1 & -2 & -2 \\ -2 & \lambda-1 & -2 \\ -2 & -2 & \lambda-1 \end{vmatrix}=\lambda^3-3\lambda^2-9\lambda-5.$$

其次，求B的特征根：

$$\lambda^3-3\lambda^2-9\lambda-5=(\lambda+1)^2(\lambda-5).$$

所以B的特征根为-1，5，当中的-1为二重根，5为单根．

再次，分别求出属于-1和5的特征向量：

1）属于-1的特征向量是齐次线性方程组

$$\begin{pmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

的非零解．显而易见这个方程组系数阵的秩为1，因而解空间的维数为2．经计算可求出由两个非零解组成的基础解系如下：

$$\alpha_1=(1,0,\ -1),\quad \alpha_2=(0,1,\ -1).$$

于是属于特征根 - 1 的特征向量可表为

$$k_1\alpha_1+k_2\alpha_2=(k_1,k_2,-k_1-k_2),\quad k_1\text{或}k_2\neq 0.$$

2）属于 5 的特征向量为齐次线性方程组

$$\begin{pmatrix}4 & -2 & -2\\ -2 & 4 & -2\\ -2 & -2 & 4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

的非零解．经计算系数阵的秩为 2，所以解空间是一维的．这样只需求出一个非零解便可表出属于 5 的一切特征向量．比如取一个非零解为

$$\beta=(1,1,1).$$

于是属于 5 的特征向量可表为

$$k\beta=k(1,1,1)=(k,k,k),\quad k\neq 0.$$

例 3　求 n 阶零阵、n 阶单位阵的特征根和特征向量．

解 1）先求 n 阶零阵的特征根．

令

$$A=\begin{pmatrix}0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 0\end{pmatrix}$$

为 n 阶零阵．于是它的特征多项式为

$$|\lambda I-A|=\begin{vmatrix}\lambda & 0 & \cdots & 0\\ 0 & \lambda & \cdots & 0\\ \cdots & \cdots & & \cdots\\ 0 & 0 & \cdots & \lambda\end{vmatrix}=\lambda^n.$$

所以零阵的特征根全都是零．即零是 n 阶零阵的 n 重根．

再求属于特征根零的特征向量．它们是齐次线性方程组

$$\begin{pmatrix}0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \cdots\\ x_n\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}$$

的非零解．显然，任一非零的 n 维向量都是 n 阶零阵属于特征根零的特征向量．

2）先求 n阶单位阵 I_n 的特征根．因为

$$|\lambda I-I_n|=\begin{vmatrix}\lambda-1 & 0 & \cdots & 0\\ 0 & \lambda-1 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & \lambda-1\end{vmatrix}=(\lambda-1)^n,$$

所以单位阵 I_n 的特征根为 1，其重数为 n。

再求 I_n 属于 1 的特征向量．它们是齐次线性方程组

$$\begin{pmatrix}0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & \cdots & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}$$

的非零解．显然，任一非零的 n维向量都是I_n 的特征向量。

练 习 二

1．求下列矩阵的特征根、特征空间和特征向量

$$A_1=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix},\ A_2=\begin{pmatrix}3 & 1 & 0\\ -4 & -1 & 0\\ 4 & -8 & 2\end{pmatrix},\ A_3=\begin{pmatrix}2 & -1 & 2\\ 5 & -3 & 3\\ -1 & 0 & -2\end{pmatrix}$$

$$A_4=\begin{pmatrix}5 & 6 & -3\\ -1 & 0 & 1\\ 1 & 2 & -1\end{pmatrix},\ A_5=\begin{pmatrix}1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1\end{pmatrix}.$$

2．我们已知零阵的特征根全为零．单位阵 I_n 的特征根全为 1．试问：特征根全为零的阵必是零阵吗？特征根全为 1 的阵必是单位阵吗？为什么？

3．我们已知对 n阶零阵、n阶单位阵来说，每一非零 n 维向量都是特征向量．试问：除了零阵、单位阵外，还有这样的阵吗？为什么？

4．试问：对相抵、相合分类来说，特征多项式是否也是不变量？为什么？

5．试问：

1）矩阵 A，B 有不同的特征根时，能否有相同的特征向量？

2）矩阵A的一个特征向量能否属于A的两个不同特征根？

6．设A为任一 n阶方阵，λ 是A的特征根，α 为A的属于特征根 λ的特征向量．试问：能否存在方阵 B，使λ 是 B 的特征根，α 为 B 的属于 λ的特征向量？

7．设任意非零的 n 维向量都是 n 阶方阵A的特征向量．证明，A为纯量阵．

8．设α 是 n阶方阵A的属于特征根 λ的特征向量，P为可逆的 n阶方阵。证明，$P^{-1}\alpha$ 是$P^{-1}AP$的属于λ的特征向量．

9．设 λ为 n阶方阵A的特征根．证明 λ^m 是 A^m 的特征根，其中 m为正整数．

10．设A为 n阶方阵．若存在正整数 m，使 $A^m=0$，则称A为幂零的；若存在正整数 m，使 $A^m=I_n$，则称A为么幂的；若适合 $A^2=A$，则称A为幂等的；若适合 $A^2=I_n$，则称A为对合阵．证明：

1）幂零阵的特征根均为 0；

2）么幂阵的n个特征根恰为n次单位根；

3）幂等阵的特征根是 0 或 1；

4）对合阵的特征根是 1 或 -1．

11．已知可逆阵A的特征根与特征向量，试求 A^{-1} 的特征根和特征向量．

12．设 α为A的特征向量，$f(\lambda)$ 为任一多项式．证明，α 是 $f(A)$ 的特征向量．

13．设 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是A的属于同一特征根 λ的特征向量．证明，$k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m$ $(\neq 0)$ 也是A的属于 λ 的特征向量．

14．证明：A与其转置阵 A' 的特征多项式相同．

15．设A是分块对角形阵：

$$A=\begin{pmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s\end{pmatrix}.$$

证明，A的特征多项式等于 $A_1,A_2,\cdots,A_s$ 的特征多项式的乘积。因此，$A_1,A_2,\cdots,A_s$ 的所有特征根恰是A的全部特征根。

16．证明：属于不同特征根的两个特征子空间的交只含零向量。

§3 特征向量系

有了§2定义 1 与定义 2 的概念，我们可以明确表述A相似于对角形阵的条件：

设 $A=(a_{ij})$，如果存在可逆矩阵 $P=(\alpha_1\alpha_2\cdots\alpha_n)$（这里 α_i 是 P 的第 i 列元素构成的），使

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

那么必有

1）$\lambda_1,\lambda_2,\cdots,\lambda_n$是$A$的全部特征根；

2）α_i 是A的属于 λ_i 的特征向量，$(i=1,2,\cdots,n)$。

不难指出，这个条件不是充分的。问题在于从A的每一个特征根 λ_i 所求出的特征向量做成 n 阶方阵 P 时未必是可逆的。因为$\alpha_1,\alpha_2,\cdots,\alpha_n$ 做成的 P 是可逆的必要而且只要 $\alpha_1,\alpha_2,\cdots,\alpha_n$是线性无关的。由上节例 1 中的$A$已经看到，从$A$的全部特征根（1和3）求不出三个线性无关的特征向量。因此，对这个阵A来说，它肯定不能相似于对角形阵。

为了探讨用A的特征向量做成可逆矩阵的可能性，以下我们深入地分析一个方阵的特征向量的线性相关性。首先，有

命题1　设 λ_0 是 n阶方阵A的 k 重特征根。那么，A的属

于 λ_0 的特征空间$S_A(\lambda_0)$的维数不超过k，即 $\dim S_A(\lambda_0)\leqslant k$。

证明　令 $\dim S_A(\lambda_0)=s$。任取 $S_A(\lambda_0)$的一个基底，用 $n\times 1$矩阵的形式记为

$$w_1,\ w_2,\ \cdots,\ w_s.$$

把这 s 个线性无关的向量扩充成为由 $n\times 1$矩阵所组成的 n 维向量空间 V_n 的基底：

$$w_1, w_2, \cdots, w_s, w_{s+1}\cdots, w_n.$$

于是

$$\begin{aligned}
Aw_1 &= \lambda_0 w_1\\
Aw_2 &= \quad \lambda_0 w_2\\
&\ \ \vdots \qquad\quad \ddots\\
Aw_s &= \qquad\quad \lambda_0 w_s\\
Aw_{s+1} &= a_{1s+1}w_1+\cdots+a_{ss+1}w_s+a_{s+1\,s+1}w_{s+1}+\cdots+a_{ns+1}w_n\\
&\ \ \vdots\\
Aw_n &= a_{1n}w_1+\cdots+a_{sn}w_s+a_{s+1\,n}w_{s+1}+\cdots+a_{nn}w_n.
\end{aligned}$$

这 n 个式子都是矩阵的等式，等号两边都是 $n\times 1$矩阵。现在把这 n 个式子集中拼写成一个矩阵的等式如下：

$$A(w_1\cdots w_s w_{s+1}\cdots w_n)=(w_1\cdots w_s w_{s+1}\cdots w_n)\left(\begin{array}{ccc|ccc}
\lambda_0 & & & a_{1s+1} & \cdots & a_{1n}\\
 & \lambda_0 & & \vdots & & \vdots\\
 & & \ddots\ \lambda_0 & a_{ss+1} & \cdots & a_{sn}\\
\hline
 & & & a_{s+1s+1} & \cdots & a_{s+1n}\\
 & 0 & & \vdots & & \cdots\\
 & & & a_{ns+1} & \cdots & a_{nn}
\end{array}\right).$$

由于 $w_1,\cdots,w_s, w_{s+1},\cdots,w_n$ 线性无关，所以它们构成的矩阵

$$W=(w_1\cdots w_s w_{s+1}\cdots w_n)$$

是一个可逆的n 阶方阵。从而上面的式子又可以写成

$$W^{-1}AW=\left(\begin{array}{cc|ccc}
\lambda_0 & & a_{1s+1} & \cdots & a_{1n}\\
 & \ddots & \vdots & & \\
 & \lambda_0 & a_{ss+1} & \cdots & a_{sn}\\
\hline
 & & a_{s+1s+1} & \cdots & a_{s+1n}\\
 0 & & \vdots & & \\
 & & a_{ns+1} & \cdots & a_{nn}
\end{array}\right).$$

这时可以看出A的特征多项式

$$|\lambda I - A| = (\lambda - \lambda_0)^s g(\lambda).$$

因为 λ_0 是 k重根，所以 $s \leqslant k$，即 $\dim S_A(\lambda) \leqslant k$，证完。

进而还有

命题2 设 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 是A的 r个不同的特征根，$\alpha_1, \alpha_2, \cdots, \alpha_r$ 分别是A的属于这 r个不同特征根的特征向量，那么$\alpha_1, \alpha_2, \cdots, \alpha_r$ 是线性无关的.

证明 对 r 用数学归纳法。首先 $r=1$时，结论显然是对的。其次假设对 $r-1$的情形结论成立，往证对 r 的情形结论也成立。为此考察线性组合

$$k_1\alpha_1 + \cdots + k_{r-1}\alpha_{r-1} + k_r\alpha_r = 0. \tag{1}$$

用A去左乘（1）式两端，则有

$$k_1(A\alpha_1) + \cdots + k_{r-1}(A\alpha_{r-1}) + k_r(A\alpha_r) = 0$$

因 α_i 是属于 λ_i 的特征向量（$i=1, \cdots, r-1, r$），所以

$$k_1\lambda_1\alpha_1 + \cdots + k_{r-1}\lambda_{r-1}\alpha_{r-1} + k_r\lambda_r\alpha_r = 0$$

再用 λ_r 去乘（1）式两端，则有

$$k_1\lambda_r\alpha_1 + \cdots + k_{r-1}\lambda_r\alpha_{r-1} + k_r\lambda_r\alpha_r = 0 \tag{2}$$

由（2）式两端分别减去（1）式两端，有

$$k_1(\lambda_r - \lambda_1)\alpha_1 + \cdots + k_{r-1}(\lambda_r - \lambda_{r-1})\alpha_{r-1} = 0$$

据归纳假设，由上式可得

$$k_1(\lambda_r - \lambda_1) = 0, \cdots, \ k_{r-1}(\lambda_r - \lambda_{r-1}) = 0$$

又因 $\lambda_1, \cdots \lambda_{r-1}, \lambda_r$ 是不同的特征根，故

$$\lambda_r - \lambda_1 \neq 0, \cdots, \lambda_r - \lambda_{r-1} \neq 0.$$

于是，必有

$$k_1 = 0, \cdots, \ k_{r-1} = 0.$$

现在回头来看（1）式，又有

$$k_r\alpha_r = 0,$$

因 α_r 是特征向量，所以是非零向量，从而可得 $k_r = 0$。总之，得到了

$$k_1=0,\cdots,\ k_{r-1}=0,\ k_r=0.$$

故 $\alpha_1,\cdots,\alpha_{r-1},\alpha_r$ 是线性无关的。

命题 3 设 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 是$\mathscr{A}$的全部不同的特征根。令

$$\alpha_{11},\alpha_{12},\cdots,\alpha_{1l_1},$$
$$\alpha_{21},\alpha_{22},\cdots,\alpha_{2l_2},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_{t1},\alpha_{t2},\cdots,\alpha_{tl_t}$$

依次为在 $S_{\mathscr{A}}(\lambda_1),S_{\mathscr{A}}(\lambda_2),\cdots,S_{\mathscr{A}}(\lambda_t)$ 中任意取出的基底，那么

$$\alpha_{11},\cdots,\alpha_{1l_1},\ \alpha_{21},\cdots,\alpha_{2l_2},\cdots,\alpha_{t1},\cdots,\alpha_{tl_t} \tag{3}$$

是线性无关的。

证明 设

$$c_{11}\alpha_{11}+\cdots+c_{1l_1}\alpha_{1l_1}+c_{21}\alpha_{21}+\cdots+c_{2l_2}\alpha_{2l_2}+\cdots+c_{t1}\alpha_{t1}+\cdots+c_{tl_t}\alpha_{tl_t}$$

$=0$，往证每个系数 c_{ij} 必全为零。

令

$$c_{11}\alpha_{11}+\cdots+c_{1l_1}\alpha_{1l_1}=u_1$$
$$c_{21}\alpha_{21}+\cdots+c_{2l_2}\alpha_{2l_2}=u_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$c_{t1}\alpha_{t1}+\cdots+c_{tl_t}\alpha_{tl_t}=u_t,$$

则有

$$u_1+u_2+\cdots+u_t=0.$$

由命题 2 必有

$$u_1=0,u_2=0,\cdots,u_t=0.$$

进而便有

$$c_{11}=0,\cdots,c_{1l_1}=0,\ c_{21}=0,\cdots,c_{2l_2}=0,\cdots,c_{t1}=0,\cdots,c_{tl_t}=0.$$

故

$$\alpha_{11},\cdots,\alpha_{1l_1},\alpha_{21},\cdots,\alpha_{2l_2},\cdots,\alpha_{t1},\cdots,\alpha_{tl_t}$$

是线性无关的。本命题证完。

从向量组（3）的取法我们断定：该向量组所含向量的个数是由$\mathscr{A}$唯一决定的，就是$\mathscr{A}$的各个特征空间维数之和，不因我们在各

个特征空间取基底时的取法不同而不同。

为了说话方便，我们给出

定义1 设A为 n阶方阵，$\lambda_1, \lambda_2, \cdots, \lambda_t$ 为A的一切不同的特征根。令

$$\alpha_{11}\cdots, \alpha_{1l_1}; \alpha_{21}, \cdots, \alpha_{2l_2}; \cdots, \quad \alpha_{t1}, \cdots, \alpha_{tl_t} \tag{4}$$

分别是在 $S_A(\lambda_1), S_A(\lambda_2), \cdots, S_A(\lambda_t)$ 中任意取出的基底。称向量组(4)为A的一个特征向量系。

显然，A的特征向量系不止一个，但所含向量的个数是由A唯一决定。又由命题1可知：若 λ_i 的重数为 $n_i (i=1,2,\cdots,t)$，则

$$l_1 \leqslant n_1, \quad l_2 \leqslant n_2, \cdots, l_t \leqslant n_t.$$

但由于 $n_1+n_2+\cdots+n_t \leqslant n$，所以必有 $l_1+l_2+\cdots+l_t \leqslant n$。即 n 阶方阵A的特征向量系中所含向量的个数不超过 n。如果 n阶方阵A的特征向量系中恰含n 个向量，则称这个特征向量系是完全的，也称A有完全的特征向量系。

例如，对上节例1中的三阶方阵

$$A = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 4 & 2 \\ 1 & -3 & 1 \end{pmatrix}$$

来说，$\alpha=(3,1,-3)$，$\beta=(-1,-1,1)$就是一个特征向量系，显然这个特征向量系是不完全的，于是A没有完全的特征向量系。

对上节例2的三阶方阵

$$B = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

来说，$\alpha_1=(1,0,-1), \alpha_2=(0,1,-1), \alpha_3=(1,1,1)$就是一个完全的特征向量系，因此B有完全的特征向量系。

由上节例3可知，n 阶零阵，n 阶单位阵都有完全的特征向量系，并且任意 n 个线性无关的向量都组成一个完全的特征向量系。

把以上的讨论概括起来，我们证明了

定理 n 阶方阵A相似于对角形阵的充分必要条件是A有完全的特征向量系，并且对任一完全特征向量系 $\alpha_1, \alpha_2, \cdots, \alpha_n$，都有

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

其中 $P = (\alpha_1\alpha_2\cdots\alpha_n)$，$\lambda_1, \lambda_2, \cdots, \lambda_n$ 为A的全部特征根。

定理不仅表述了一个方阵与对角形阵相似的条件，同时也给出了演化矩阵的求法。即 $P = (\alpha_1\alpha_2\cdots\alpha_n)$ 是使A相似于对角形阵的演化矩阵的充分必要条件：$\alpha_1, \alpha_2, \cdots, \alpha_n$是$A$的一个完全特征向量系。而完全特征向量系就由各个特征空间的基底组成，也就是由各个特征方程组的基础解系组成的。

例如，对三阶方阵

$$B = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

来说，它有完全的特征向量系：

$$\alpha_1 = (1, 0, -1), \alpha_2 = (0, 1, -1), \alpha_3 = (1, 1, 1)。$$

于是

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$$

就是一个演化矩阵，并且有

$$P^{-1}AP = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 5 \end{pmatrix}.$$

下面证明特征多项式的一个重要性质，来结束这一节的讨论。

哈米顿-凯莱（Hamilton-Cayley）定理 设 n 阶方阵A的特征多项式为

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

那么

$$f(A)=A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nA^0=0.$$

即方阵A是其特征多项式的根，其中 $A^0=I_n$。

证明　设 $B(\lambda)$是 $\lambda I-A$ 的伴随矩阵，则有

$$B(\lambda)(\lambda I-A)=f(\lambda)I_n \qquad (5)$$

因为 $B(\lambda)$ 的元素是$\lambda I-A$ 中元素的代数余子式，都是 λ 的多项式，次数不超过$n-1$。因此，按矩阵运算的规则，$B(\lambda)$ 可以写成以下形式：

$$B(\lambda)=B_0\lambda^{n-1}+B_1\lambda^{n-2}+\cdots+B_{n-2}\lambda+B_{n-1}$$

其中 $B_0,B_1,\cdots,B_{n-2}$, B_{n-1} 都是 n 阶方阵。

于是

$$B(\lambda)(\lambda I-A)=(B_0\lambda^{n-1}+B_1\lambda^{n-2}+\cdots+B_{n-2}\lambda+B_{n-1})(\lambda I-A)$$
$$=B_0\lambda^n+(B_1-B_0A)\lambda^{n-1}+\cdots+(B_{n-1}-B_{n-2}A)\lambda-B_{n-1}A.$$

而

$$f(\lambda)I_n=I_n\lambda^n+a_1I_n\lambda^{n-1}+\cdots+a_{n-1}I_n\lambda+a_nI_n.$$

由（5）式比较两端λ的同次项的系数，便得

$$\begin{aligned}B_0&=I_n\\B_1-B_0A&=a_1I_n\\B_2-B_1A&=a_2I_n\\&\vdots\\B_{n-1}-B_{n-2}A&=a_{n-1}I_n\\-B_{n-1}A&=a_nI_n.\end{aligned}$$

以 A^n, A^{n-1}, $\cdots$, A, $A^0=I_n$ 依次从右边去乘上式两端，得

$$\begin{aligned}B_0A^n&=A^n\\B_1A^{n-1}-B_0A^n&=a_1A^{n-1}\\B_2A^{n-2}-B_1A^{n-1}&=a_2A^{n-2}\\&\vdots\\B_{n-1}A-B_{n-2}A^2&=a_{n-1}A\\-B_{n-1}A&=a_nI_n.\end{aligned}$$

把上述 $n+1$ 个式子两端一起相加，左端等于零，右端恰为

$$A^n + a_1A^{n-1} + a_2A^{n-2} + \cdots + a_{n-1}A + a_nA^0$$

故有

$$f(A) = A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nA^0 = 0.$$

本定理证完。

由哈米顿——凯莱定理知，对任意 n 阶方阵 A，总有非零多项式 $g(\lambda)$，使 $g(A) = 0$，称这样的多项式为 A 的化零多项式。比如 A 的特征多项式就是 A 的一个化零多项式。因此，A 的一切化零多项式中，必有非零的且次数为最低的。我们给出

定义 2　设 n 阶方阵 A。A 的化零多项式中，非零的且又次数最低、首项系数为 1 的多项式，叫做 A 的最小多项式。

例如，纯量阵

$$A = \begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{pmatrix}$$

的化零多项式中，非零的次数最低的而又首项系数为 1 的多项式为 $\lambda - a$，于是 $\lambda - a$ 就是纯量阵 A 的最小多项式。

又如，阵

$$B = \begin{pmatrix} 1 & & \\ & 2 & \\ & & 3 \end{pmatrix}$$

的特征多项式为 $(\lambda-1)(\lambda-2)(\lambda-3)$，且 B 的化零多项式的次数不能再低了，故 B 的最小多项式也为 $(\lambda-1)(\lambda-2)(\lambda-3)$。

再如，阵

$$C = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}$$

的特征多项式为 $(\lambda-2)(\lambda-3)^2$，而 C 的化零多项式的次数不可能再低了，故 C 的最小多项式也为 $(\lambda-2)(\lambda-3)^2$。

命题 4　设 $\varphi(\lambda)$ 为 A 的最小多项式。若 $g(\lambda)$ 为 A 的任一化零

多项式，则$\varphi(\lambda)\mid g(\lambda)$。

证明　按带余除法，可写出

$$g(\lambda)=\varphi(\lambda)q(\lambda)+r(\lambda)$$

其中 $r(\lambda)=0$ 或者$\deg r(\lambda)<\deg\varphi(\lambda)$。

于是，用A代替λ，便有

$$g(A)=\varphi(A)q(A)+r(A)$$

因为 $\varphi(\lambda)$、$g(\lambda)$ 都是A的化零多项式，即 $\varphi(A)=0$，$g(A)=0$，所以必有 $r(A)=0$。这样可以断定 $r(\lambda)=0$。不然 $r(\lambda)\neq0$，又$\deg r(\lambda)<\deg\varphi(\lambda)$ 且有 $r(A)=0$。这与 $\varphi(\lambda)$是A的最小多项式相矛盾。故 $\varphi(\lambda)\mid g(\lambda)$，证完。

由于一个方阵A的特征多项式是A的化零多项式，据命题 4 可知，一个方阵A的最小多项式能整除它的特征多项式。因而要求方阵 A的最小多项式，就可以到它的特征多项式的因子当中去找，找以A为根、次数最小、而又首项系数为 1 的因子就是。

例如，方阵

$$A=\begin{pmatrix}2&&&\\&2&&\\&&3&\\&&&3\end{pmatrix}$$

的特征多项式为 $(\lambda-2)^2(\lambda-3)^2$，其因子为：

$$\lambda-2,\ (\lambda-2)^2,\ \lambda-3,\ (\lambda-3)^2,\ (\lambda-2)(\lambda-3),\ (\lambda-2)^2(\lambda-3),$$
$$(\lambda-2)(\lambda-3)^2,\ (\lambda-2)^2(\lambda-3)^2.$$

从次数最低的因子找起，显然前四个不是以A为根，而

$$(A-2I_4)(A-3I_4)=\begin{pmatrix}0&&&\\&0&&\\&&1&\\&&&1\end{pmatrix}\begin{pmatrix}-1&&&\\&-1&&\\&&0&\\&&&0\end{pmatrix}=0$$

故A的最小多项式为 $(\lambda-2)(\lambda-3)$。

练 习 三

1．求上节练习的第1题各阵的特征向量系，进一步指出那些方阵相似于对角形阵，且写出其对角形阵及演化矩阵．

2．n 阶方阵A相似于对角形阵的充分必要条件是A有 n 个线性无关的特征向量．这个结论对吗？为什么？

3　如果 n阶方阵A有特征根 λ_0, $S_A(\lambda_0)$的维数小于 λ_0 的重数称A为亏损的，否则称A为非亏损的．求证：n 阶方阵A是非亏损的充分必要条件A有 n 个线性无关的特征向量．

4．证明：若A的特征根都是单根时，则A与对角形阵相似；反之，若A与对角形阵相似，那么A的特征根都是单根，对吗？为什么？

5．证明：n 阶方阵A与对角形阵相似的充分必要条件是任何s重根 λ，$\lambda I-A$的秩等于 $n-s$．

6．设

$$A=\begin{pmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{pmatrix}$$

求 A^k．$(k>0)$．

7　设 α,β 是A的属于不同的特征根的特征向量，则当$kl\neq 0$时，$k\alpha+l\beta$不是A的特征向量．

8．设A是一个 n 阶下三角形阵．证明：

1）若 $a_{ii}\neq a_{jj}(i\neq j)$，$i,j=1,2,\cdots,n$，那么$A$相似于对角形阵；

2）若$a_{ii}=a_{jj}$，$i,j=1,2,\cdots,n$．而至少有一个 $a_{i_0j_0}\neq 0$ $(i_0>j_0)$，那么A不与对角形阵相似．

9．求下列各阵的最小多项式：

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 3 & \\ & & & 3 \end{pmatrix}.$$

10 证明：A与A'的最小多项式相同。

11．试问：当两个n阶方阵的特征多项式相同，那么最小多项式也必相同吗？为什么？

12．设对角形分块阵

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix},$$

则A的最小多项式恰是$A_1, A_2, \cdots, A_s$的最小多项式的最小公倍式。

§4 正交矩阵

上一章我们证明了对称阵都相合于对角形阵。前一节一般地讨论了方阵A相似于对角形阵的条件。以下两节将证明：实对称阵都相似于对角形阵。而且能够“一举两得”，即证明存在实可逆阵P，满足条件

$$P' = P^{-1} \tag{1}$$

且使

$$P'AP = P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}. \tag{2}$$

满足条件（1）的可逆矩阵P及对于实对称阵A存在这样的P，使（2）成立，二者都有明确的几何意义和重要应用。为了能想得清楚些，我们还是从大家都熟悉的二次曲线化简说起。

在解析几何里，化简中心在坐标原点的二次曲线

$$ax^2 + 2bxy + cy^2 = d$$

使其只含坐标的平方项：

$$a'x'^2+c'y'^2=d.$$

所用的坐标变换是

$$x=x'\cos\theta-y'\sin\theta$$
$$y=x'\sin\theta+y'\cos\theta.$$

此处坐标变换的系数阵是

$$P=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

上述事实用矩阵的形式来表达就是：对二阶实对称阵

$$A=\begin{pmatrix}a & b\\ b & c\end{pmatrix}.$$

存在实可逆矩阵

$$P=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

使

$$P'AP=\begin{pmatrix}a' & 0\\ 0 & c'\end{pmatrix}.$$

我们注意到，这里的演化矩阵 P 不是一般的可逆矩阵，它具有很特殊的性质，就是

$$P'=P^{-1}.$$

因此，以这样的 P 为演化矩阵时，不仅使 A 相合于一个对角形阵，同时也使 A 相似于这个对角形阵．

对于空间二次曲面的化简问题也有类似的情况．由此推而广之，我们就自然的这样提出问题：对于一般的 n 阶实对称阵 A，是否存在满足条件

$$P'=P^{-1}$$

的可逆阵 P 使

$$P'AP=P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

这就是以下两节所要研究解决的问题．在这一节里先讨论满足条件（1）的矩阵的一般性质以及这样的矩阵的具体构造方法．有了这些准备，在下一节再来证明主要结果：对任一实对称阵A，存在满足条件 $P'=P^{-1}$ 的可逆阵P，使

$$P'AP=P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

定义1 设T为实 n 阶方阵．如果

$$T'=T^{-1},$$

则称T为正交矩阵．

这样以正交矩阵T为演化矩阵，作用于矩阵A时，得到的矩阵B：

$$T'AT=B=T^{-1}AT$$

既相合于A又相似于A．以下称这样的B是与A正交相合的或正交相似的（意思是：正交相合＝既相合又相似；正交相似＝既相似又相合）．

例如

$$A=\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix},\quad B=\begin{pmatrix}-1 & & \\ & -1 & \\ & & 1\end{pmatrix},\quad C=\begin{pmatrix}\frac{2}{3} & \frac{2}{3} & -\frac{1}{3}\\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3}\\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\end{pmatrix}$$

都是正交阵．

事实上，$A'=\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix}=A^{-1}$，$B'=\begin{pmatrix}-1 & & \\ & -1 & \\ & & 1\end{pmatrix}=B^{-1}$，$C'=C=C^{-1}$．

下面给出正交矩阵的一些基本性质．

命题1 若A为正交矩阵，则A',A^{-1}，A^*（A的伴随矩阵）都是正交矩阵；若 A,B 都是正交矩阵，则 AB也是正交矩阵．

证明 由A是正交阵，即 $A'=A^{-1}$，于是

$(A')'=A$，$(A')^{-1}=(A^{-1})^{-1}=A$，即 $(A')'=(A')^{-1}$，$(A^{-1})'=(A^{-1})^{-1}$。所以 A'，A^{-1} 是正交矩阵。而

$(A^*)'=(A')^*=(A^{-1})^*=(A^*)^{-1}$，即$(A^*)'=(A^*)^{-1}$。所以$A^*$是正交矩阵。

又 $(AB)'=B'A'=B^{-1}A^{-1}=(AB)^{-1}$，即 $(AB)'=(AB)^{-1}$，故 AB 也是正交阵。命题证完。

命题2 若$\boldsymbol{A}$是正交矩阵，则

1）$\boldsymbol{A}$的行列式等于 1 或 -1；

2）$\boldsymbol{A}$的特征根的模等于 1。

证明 1）由 $A'=A^{-1}$，$AA'=I$。于是 $|A||A'|=1$，即 $|A|^2=1$，所以 $|A|=\pm 1$；

2）设 λ 为A的任一特征根，于是有非零的 $n\times 1$矩阵 α，使

$$A\alpha=\lambda\alpha. \tag{1}$$

两端取转置，则有

$$\alpha' A'=\lambda\alpha'$$

（1）式两端取共轭，则有

$$\overline{A}\,\overline{\alpha}=\overline{\lambda}\,\overline{\alpha}$$

从而可得

$$(\alpha' A')(\overline{A}\,\overline{\alpha})=(\lambda\alpha')(\overline{\lambda}\,\overline{\alpha})=\lambda\overline{\lambda}\cdot\alpha'\overline{\alpha}.$$

因为 $A'=A^{-1}$，$\overline{A}=A$，故有

$$\alpha'\overline{\alpha}=\lambda\overline{\lambda}\cdot\alpha'\overline{\alpha} \text{ 或 } (\lambda\overline{\lambda}-1)\alpha'\overline{\alpha}=0$$

因为α是非零的，所以 $\alpha'\overline{\alpha}\neq 0$。这样便得

$\lambda\overline{\lambda}-1=0$，或 $\lambda\overline{\lambda}=1$，即 $|\lambda|=1$，证完。

命题3 下列三条是等价的：

1）实方阵 $\boldsymbol{A}=(a_{ij})$是正交矩阵；

2）$\boldsymbol{AA'=I}$ $(\boldsymbol{A'A=I})$；

3) $a_{i1}a_{j1}+a_{i2}a_{j2}+\cdots+a_{in}a_{jn}=\begin{cases}1, 当\ i=j\ 时\\0, 当\ i\neq j\ 时\end{cases}$

$$\left(a_{1i}a_{1j}+a_{2i}a_{2j}+\cdots+a_{ni}a_{nj}=\begin{cases}1, 当\ i=j 时\\0\ 当\ i\neq j 时\end{cases}\right).$$

事实上，由定义直接从 1）得到 2）．而 2）写成元素表达形式就得到 3）再据定义由 3）可得到 1）．证完．

例 1　设 $A=\begin{pmatrix}A_1 & \\ & A_2\end{pmatrix}$，其中 A_1, A_2 都是方的．那么 A 是正交矩阵的充分必要条件 A_1, A_2 都是正交矩阵．

解　$A'=\begin{pmatrix}A_1' & \\ & A_2'\end{pmatrix}$，于是

$$AA'=\begin{pmatrix}A_1A_1' & \\ & A_2A_2'\end{pmatrix}.$$

从而 AA' 是单位阵的充分必要条件 A_1A_1' 与 A_2A_2' 同时都是单位阵．

例 2　设三角形阵

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ & a_{22} & \cdots & a_{2n}\\ & & \ddots & \vdots\\ 0 & & & a_{nn}\end{pmatrix}.$$

证明：若 A 为正交矩阵，则 A 必为对角形阵，即 $a_{ij}=0\ (i<j)$．

证明　对 n 用数学归纳法．$n=1$ 时，自然是对的．假设 $n-1$ 时是对的．考虑 A 是 n 阶正交矩阵的情形．

首先 $a_{11}\neq 0$．于是由命题 3 的 3）知：从第一列看应有 $a_{11}^2=1$，从第一行看应有 $a_{11}^2+a_{12}^2+\cdots+a_{1n}^2=1$，故可推得 $a_{12}=0, \cdots, a_{1n}=0$．即

$$A=\begin{pmatrix}a_{11} & 0 & \cdots & 0\\ & a_{22} & \cdots & a_{2n}\\ & & \ddots & \vdots\\ 0 & & & a_{nn}\end{pmatrix}=\begin{pmatrix}a_{11} & \\ & A_1\end{pmatrix}.$$

由例 1 可知，A_1 是 $n-1$ 阶三角形正交矩阵．这样，据归纳法假设便知 A_1 必为对角形的，从而 A 就是对角形阵．

例3 设$A=\begin{pmatrix} a & c \\ b & d \end{pmatrix}$为二阶正交矩阵．试证：

1）当 $|A|=1$时，$A=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

2）当 $|A|=-1$ 时，$A=\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$.

证明

1）当 $|A|=1$时，有

$A^{-1}=\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$，$A'=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，于是 $d=a, c=-b$．从而，有

$A=\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$且 $a^2+b^2=1$．令

$a=\cos\theta$，取 $b=\sin\theta$

则有

$$A=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

2）当 $|A|=-1$时，有

$A^{-1}=\begin{pmatrix} -d & c \\ b & -a \end{pmatrix}$，$A'=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，于是 $d=-a$，$b=c$．从而，有

$A=\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$，且 $-a^2-b^2=-1$，即 $a^2+b^2=1$．

令

$a=\cos\theta$，取 $b=\sin\theta$

则得

$A=\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$，证完．

在命题3中，用元素表述的A为正交矩阵的条件相当重要，依

据该条件可以给出构造正交矩阵的一般方法。为此，我们具体地考察一下它们的特点：

$$a_{i1}a_{j1}+a_{i2}a_{j2}+\cdots+a_{in}a_{jn}=\begin{cases}1, i=j \text{ 时}\\0, i\neq j \text{ 时}\end{cases} \quad (2)$$

此处

$$a_{i1}, a_{i2}, \cdots, a_{in} \text{ 与 } a_{j1}, a_{j2}, \cdots, a_{jn}$$

恰好是A的第 i 行与第 j 行元素。于是 (2)式说明：

A的不同两行元素对应乘积之和等于零；

A的同一行元素对应地自乘的积相加之和等于1。

完全类似地，A的各列元素也有相应的性质。

正交矩阵各行（列）元素的这种性质对我们来说并不是很陌生的。由于 n 阶方阵的行（列）都可以看做是 n 维向量，于是若令

$$\alpha_i=(a_{i1}, a_{i2}, \cdots, a_{in})$$

$$\beta'_i=(a_{1i}, a_{2i}, \cdots, a_{ni}),\quad i=1,2,\cdots,n.$$

则有

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\a_{21} & a_{22} & \cdots & a_{2n}\\\cdots & \cdots & \cdots & \cdots\\a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}=\begin{pmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{pmatrix}=(\beta_1\beta_2\cdots\beta_n)$$

这样，上述正交矩阵A各行（列）元素的性质可写成：

$$\alpha_i\alpha_j'=\begin{cases}1, i=j \text{ 时}\\0, i\neq j \text{ 时}\end{cases}$$

和

$$\beta_i'\beta_j=\begin{cases}1, i=j \text{ 时}\\0, i\neq j \text{ 时}\end{cases}$$

当 $n=2,3$时，这正是解析几何里二、三维向量的数积运算。而它们的几何意义是：

$\alpha_i\alpha_j'=0$ 必要而且只要 α_i 与 α_j 垂直也叫正交；

$\alpha_i\alpha_i'=1$ 必要而且只要 α_i 的长等于1也叫 α_i 为单位向量或标

准向量。

虽然当 $n>3$ 时，n 维向量已没有直观的几何形象了，但我们还是借用了向量这个几何名称。把 n 数组，$n\times1$ 矩阵及 $1\times n$ 矩阵都统称为 n 维向量。这样做不仅仅是一种习惯和简单模仿，而是由此常常会给我们造成一些有益的联想，通过这些联想去探索、研究 n 维向量的某些与 2 、 3 维向量相类似的性质和应用。正是因为这样，我们又有

定义 2 设 $\alpha=(a_1,a_2,\cdots,a_n)$，$\beta=(b_1,b_2,\cdots,b_n)$ 为任二实 n 维向量。如果

$$\alpha\beta'=a_1b_1+a_2b_2+\cdots+a_nb_n=0$$

则称 α 与 β 正交，记作 $\alpha\perp\beta$；

如果

$$\alpha\alpha'=a_1^2+a_2^2+\cdots+a_n^2=1$$

则称 α 为单位向量或标准向量，记作 $|\alpha|=1$。

定义 3 设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 为一组 n 维向量。

1）如果 $|\alpha_i|=1$，$(i=1,2,\cdots,s)$，则称 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 为一标准向量组；

2）对任意的 α_i,α_j，如果 $\alpha_i\perp\alpha_j\,(i\neq j)$，则称 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 为一正交向量组，特别的，当 $s=1$ 时也认为 α_1 是一正交向量组；

3）如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是一标准向量组又是正交向量组，则称其为标准正交向量组；

4）设 β 是一 n 维向量，如果向量 β 与每一 $\alpha_i\,(i=1,2,\cdots,s)$ 都正交，即

$$\beta\perp\alpha_i,\quad i=1,2,\cdots,s.$$

就说 β 与向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 正交。两组向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 与 $\beta_1,\beta_2,\cdots,\beta_r$（是一组 n 维向量），如果每一 α_i 与每一 β_j 都正交，则称这两个向量组是正交的。

这里要注意的是：两组向量正交与这两组向量本身是否是正交的毫无关系。

按着这一定义，有

A是正交矩阵必要而且只要A的行向量是一标准正交组；A是正交矩阵必要而且只要A的列向量是一标准正交组。

又比如

$\alpha_1=\left(\frac{\sqrt{2}}{2},\ \frac{\sqrt{2}}{2},\ 0\right),\alpha_2=\left(\frac{\sqrt{2}}{2},\ 0,\ \frac{\sqrt{2}}{2}\right)$为标准向量组；而

$\beta_1=(1,0,0)$，$\beta_2=(0,2,0)$，$\beta_3=(0,0,3)$ 为正交向量组；而

$\varepsilon_1=(1,0,0)$，$\varepsilon_2=(0,1,0)$，$\varepsilon_3=(0,0,1)$ 与

$\eta_1=(1,0,0,0)$，$\eta_2=(0,1,0,0)$，$\eta_3=(0,0,1,\ 0)$，$\eta_4=(0,0,0,1)$ 都是标准正交向量组；

向量 $\gamma=(2,0,0,0)$ 与向量组 $\delta_1=(0,1,0,0)$，$\delta_2=(0,0,2,0)$，$\delta_3=(0,0,0,3)$ 正交；

向量组 $\alpha_1=(-1,0,0,0)$，$\alpha_2=(0,-1,0,0)$ 与向量组 $\beta_1=(0,0,-2,0)$，$\beta_2=(0,0,0,-2)$，$\beta_3=(0,0,0,-1)$ 是正交的。

命题 4　非零正交向量组必线性无关。

证明　设非零正交向量组$\alpha_1,\alpha_2,\cdots,\alpha_s$. 考察

$$k_1\alpha_1+\cdots+k_i\alpha_i+\cdots+k_s\alpha_s=0.$$

一方面

$$(k_1\alpha_1+\cdots+k_i\alpha_i+\cdots+k_s\alpha_s)\alpha_i'=0.$$

另一方面

$$\begin{aligned}&(k_1\alpha_1+\cdots+k_i\alpha_i+\cdots+k_s\alpha_s)\alpha_i'\\&=k_1\alpha_1\alpha_i'+\cdots+k_i\alpha_i\alpha_i'+\cdots+k_s\alpha_s\alpha_i'\\&=k_i\alpha_i\alpha_i'.\end{aligned}$$

于是便得

$$k_i\alpha_i\alpha_i'=0$$

因 $\alpha_i\neq0$，所以 $\alpha_i\alpha_i'\neq0$，因而 $k_i=0$，$i=1,2,\cdots,s$. 故$\alpha_1,\alpha_2,\cdots,\alpha_s$ 是线性无关的，证完。

命题 5　设以向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 为行的矩阵是

$$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_s\end{pmatrix}.$$

那么，向量 $\beta=(b_1,b_2,\cdots,b_n)$ 与向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 正交的充分必要条件：β是齐次线性方程组

$$AX=0$$

的解。

证明　令

$$\alpha_1=(a_{11},a_{12},\cdots,a_{1n})$$
$$\alpha_2=(a_{21},a_{22},\cdots,a_{2n})$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\alpha_s=(a_{s1},a_{s2},\cdots,a_{sn}).$$

于是

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{s1} & a_{s2} & \cdots & a_{sn}\end{pmatrix}.$$

以A为系数阵的齐次线性方程组就是

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{s1}x_1+a_{s2}x_2+\cdots+a_{sn}x_n=0.\end{cases}$$

这样 $\beta=(b_1,b_2,\cdots,b_n)$ 是上述齐次线性方程组的解充分必要条件是

$$\begin{cases}a_{11}b_1+a_{12}b_2+\cdots+a_{1n}b_n=0\\ a_{21}b_1+a_{22}b_2+\cdots+a_{2n}b_n=0\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{s1}b_1+a_{s2}b_2+\cdots+a_{sn}b_n=0.\end{cases}\qquad(3)$$

（3）成立的充分必要条件是：

$$\alpha_1\beta'=0,\ \alpha_2\beta'=0,\cdots,\ \alpha_s\beta'=0$$

即

$$\alpha_1\perp\beta,\ \alpha_2\perp\beta,\cdots,\ \alpha_s\perp\beta.$$ 证完。

命题5虽然简单但很重要，它说明 n维向量的正交性与齐次线性方程组的密切联系. 特别地，它给出求与已知一组向量正交的向量的一般方法，从而也解决了构造一组正交向量的一般方法，这全部都归结为解齐次线性方程组.

下面就来给出构造正交矩阵的方法.

定理1 设 $\beta_1,\beta_2,\cdots,\beta_r$ 是任意 r 个线性无关的n 维向量. 那么按下列方式都能得到唯一的一个标准正交向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$：

$$
\begin{aligned}
&\alpha_1=b_{11}\beta_1\\
&\alpha_2=b_{21}\beta_1+b_{22}\beta_2\\
&\quad\vdots\\
&\alpha_r=b_{r1}\beta_1+b_{r2}\beta_2+\cdots+b_{rr-1}\beta_{r-1}+b_{rr}\beta_r
\end{aligned}
$$

其中 $b_{11},b_{22},\cdots,b_{rr}$都是正数.

证明　我们的证法是具体地一个一个的构造出 $\alpha_1,\alpha_2,\cdots,\alpha_r$. 这只要依据标准、正交这两条要求来决定系数 b_{ij} 就行了.

首先构造 α_1.

α_1 应是标准向量，即 $|\alpha_1|=1$. 于是 $\alpha_1\alpha_1'=b_{11}\beta_1\cdot b_{11}\beta_1'=b_{11}^2\beta_1\beta_1'$. 而$\beta_1\neq0$，由此则有

$$b_{11}^2=\frac{1}{\beta_1\beta_1'}，\text{所以取}b_{11}=\frac{1}{\sqrt{\beta_1\beta_1'}}，\text{就得}$$

$$\alpha_1=b_{11}\beta_1，\ |\alpha_1|=1.$$

其次构造 α_2.

α_2 应与 α_1 正交，并且本身是标准向量. 为此我们先令

$\bar{\alpha}_2=c_{21}\alpha_1+\beta_2$，使 $\bar{\alpha}_2\perp\alpha_1$，由此确定$\bar{\alpha}_2$，然后再把 $\bar{\alpha}_2$ 标准化为α_2，这样得到的α_2即合所求.

事实上，$\bar{\alpha}_2\perp\alpha_1$ 必要而且只要$\bar{\alpha}_2\alpha_1'=0$，即

$$(c_{21}\alpha_1+\beta_2)\alpha_1'=c_{21}\alpha_1\alpha_1'+\beta_2\alpha_1'=c_{21}+\beta_2\alpha_1'=0.$$

于是

$c_{21}=-\beta_2\alpha_1'$，故 $\bar{\alpha}_2$ 就可求出.

又因$\bar{\alpha}_2\neq0$. 不然若$\bar{\alpha}_2=0$，即 $c_{21}\alpha_1+\beta_2=c_{21}b_{11}\beta_1+\beta_2=0$ 这与 β_1,β_2

是线性无关的矛盾. 故取

$$\alpha_2=\frac{\tilde{\alpha}_2}{\sqrt{\tilde{\alpha}_2\tilde{\alpha}_2'}}=-\frac{\beta_2\alpha_1'}{\sqrt{\tilde{\alpha}_2\tilde{\alpha}_2'}}\alpha_1+\frac{1}{\sqrt{\tilde{\alpha}_2\tilde{\alpha}_2'}}\beta_2$$

$$=b_{21}\beta_1+b_{22}\beta_2$$

其中

$$b_{21}=-\frac{\beta_2\alpha_1'}{\sqrt{\tilde{\alpha}_2\tilde{\alpha}_2'}\sqrt{\beta_1\beta_1'}},\quad b_{22}=\frac{1}{\sqrt{\tilde{\alpha}_2\tilde{\alpha}_2'}}.$$

都由 β_1,β_2 所唯一决定, 且 b_{22} 是正数.

再其次构造 α_3. 为了简单, 我们先考虑 $\alpha_1,\alpha_2,\beta_3$ 的一个线性组合

$$\tilde{\alpha}_3=c_{31}\alpha_1+c_{32}\alpha_2+\beta_3$$

使其与已做出的 α_1,α_2 都正交, 由此确定系数 c_{31} 与 c_{32}. 而由于 β_1,β_2,β_3 线性无关, 所以 $\tilde{\alpha}_3\neq 0$. 再把$\tilde{\alpha}_3$标准化为单位向量 α_3, 那么 α_3 即合所求.

事实上, $\tilde{\alpha}_3\perp\alpha_1$: $\tilde{\alpha}_3\alpha_1'=(c_{31}\alpha_1+c_{32}\alpha_2+\beta_3)\alpha_1'=c_{31}+\beta_3\alpha_1'=0$,

$\tilde{\alpha}_3\perp\alpha_2$: $\tilde{\alpha}_3\alpha_2'=(c_{31}\alpha_1+c_{32}\alpha_2+\beta_3)\alpha_2'=c_{32}+\beta_3\alpha_2'=0$.

于是, $c_{31}=-\beta_3\alpha_1',c_{32}=-\beta_3\alpha_2'$, $\tilde{\alpha}_3=-(\beta_3\alpha_1')\alpha_1-(\beta_3\alpha_2')\alpha_2+\beta_3$.

取

$$\alpha_3=\frac{\tilde{\alpha}_3}{\sqrt{\tilde{\alpha}_3\tilde{\alpha}_3'}}=-\frac{\beta_3\alpha_1'}{\sqrt{\tilde{\alpha}_3\tilde{\alpha}_3'}}\alpha_1-\frac{\beta_3\alpha_2'}{\sqrt{\tilde{\alpha}_3\tilde{\alpha}_3'}}\alpha_2+\frac{1}{\sqrt{\tilde{\alpha}_3\tilde{\alpha}_3'}}\beta_3$$

$$=b_{31}\beta_1+b_{32}\beta_2+b_{33}\beta_3.$$

系数 b_{31},b_{32},b_{33} 都由 β_1,β_2,β_3 所唯一决定且b_{33}是正数. 则$\alpha_1,\alpha_2,\alpha_3$是一标准正交组.

按照这样的程序做下去, 自然就得到定理 1 中所说的那样一个标准正交组. 证完.

定理 1 的证明过程就是从一组给定的线性无关向量 $\beta_1,\beta_2,\cdots,\beta_r$ 构造出一个标准正交组 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 的具体做法, 把这种方法叫做 Schmidt 正交化法. 这样, 用 Schmidt正交化法从任意一组线性无关向量 $\beta_1,\beta_2,\cdots,\beta_r$都能求出唯一的一个标准正交组 $\alpha_1,\alpha_2,\cdots,$

α_r，它们被 $\beta_1,\beta_2,\cdots,\beta_r$线性表出如定理 1 所说的形式。

定理 1 还有另外两种表述方式：矩阵方式和空间方式。我们分别列出如下：

定理1′　设 $A=(a_{ij})$为 $r\times n$ 矩阵，且$\text{rank}A=r$。则存在 r 阶下三角形阵P，使

$$B=PA$$

的行向量为一标准正交组。

推论 1　若 $A=(a_{ij})$是 $n\times r$矩阵，且$\text{rank}\,A=r$。则存在r阶上三角形阵Q，使

$$B=AQ$$

的列向量为一标准正交组。

推论 2　对任一可逆阵A，存在下（上）三角形阵 $P(Q)$，使

$$T_1=PA\quad(T_2=AQ)$$

为正交矩阵。

定理1″　设V是 $n(n\geqslant1)$维向量空间，那么V存在由标准正交组构成的基底。

定理 2　设 $T_0=(t_{ij})$为 $r\times n$ 矩阵，它的行向量是一个标准正交组。那么，存在 $(n-r)\times n$ 矩阵 T_1，使 $T=\begin{pmatrix}T_0\\T_1\end{pmatrix}$成为 n阶正交矩阵。

证明　显然$\text{rank}\,T_0=r$。考察以T_0为系数阵的齐次线性方程组

$$T_0X=0.$$

于是它的解空间W是$n-r$维的。由定理1″，任取W的由标准正交组构成的基底，把它做成行矩阵 T_1。这样

$$T=\begin{pmatrix}T_0\\T_1\end{pmatrix}$$

就是一个 n阶正交矩阵。证完。

推论　设 S_0 为 $n\times r$矩阵，它的列向量是一标准正交组。那么存在 $n\times(n-r)$矩阵 S_1，使 $S=(S_0S_1)$成为 n阶正交矩阵。

定理 1 和定理 2 给出一个具体的构造正交矩阵的方法。

下面通过例题来熟悉一下构造正交矩阵的方法。

例 1　设

$$A=\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ 0 & 1 & 2\end{pmatrix}$$

求下三角形阵 P，使 PA 为正交矩阵。

解　把 A 的行做为三个向量

$$\beta_1=(1,1,0),\quad \beta_2=(1,-1,0),\quad \beta_3=(0,1,2)$$

首先，令

$$\alpha_1=\frac{1}{\sqrt{\beta_1\beta_1'}}\beta_1=\frac{1}{\sqrt{2}}(1,1,0)=\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},0\right)$$

其次，令

$$\bar{\alpha}_2=c_{21}\alpha_1+\beta_2=c_{21}\left(\frac{1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},\ \ 0\right)+(1,\ -1,\ 0)=\left(\frac{c_{21}}{\sqrt{2}}+1,\frac{c_{21}}{\sqrt{2}}-1,\ 0\right)$$

使 $\bar{\alpha}_2$ 与 α_1 正交，则有

$$\bar{\alpha}_2\alpha_1'=\left(\frac{c_{21}}{\sqrt{2}}+1,\frac{c_{21}}{\sqrt{2}}-1,\ 0\right)\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\\ 0\end{pmatrix}=\left(\frac{c_{21}}{\sqrt{2}}+1\right)\frac{1}{\sqrt{2}}$$

$$+\left(\frac{c_{21}}{\sqrt{2}}-1\right)\frac{1}{\sqrt{2}}=c_{21}=0$$

所以

$$\bar{\alpha}_2=\beta_2.$$

将 $\bar{\alpha}_2$ 标准化为 α_2，即令

$$\alpha_2=\frac{1}{\sqrt{\bar{\alpha}_2\bar{\alpha}_2'}}\bar{\alpha}_2=\frac{1}{\sqrt{2}}(1,-1,0)=\left(\frac{1}{\sqrt{2}},\ \ -\frac{1}{\sqrt{2}},\ \ 0\right),$$

最后，令

$$\tilde{\alpha}_3 = c_{31}\alpha_1 + c_{32}\alpha_2 + \beta_3$$

使 $\tilde{\alpha}_3$与 α_1，α_2都正交，则有

$$\tilde{\alpha}_3\alpha_1' = (c_{31}\alpha_1 + c_{32}\alpha_2 + \beta_3)\alpha_1' = c_{31} + \beta_3\alpha_1' = 0$$

从而

$$c_{31} = -\beta_3\alpha_1' = -(0, 1, 2)\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = -\frac{1}{\sqrt{2}}.$$

$$\tilde{\alpha}_3\alpha_2' = (c_{31}\alpha_1 + c_{32}\alpha_2 + \beta_3)\alpha_2' = c_{32} + \beta_3\alpha_2' = 0$$

从而

$$c_{32} = -\beta_3\alpha_2' = -(0, 1, 2)\begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}$$

所以

$$\tilde{\alpha}_3 = -\frac{1}{\sqrt{2}}\alpha_1 + \frac{1}{\sqrt{2}}\alpha_2 + \beta_3 = \left(-\frac{1}{2}, -\frac{1}{2}, 0\right) + \left(\frac{1}{2}, -\frac{1}{2}, 0\right)$$

$$+ (0, 1, 2) = (0, 0, 2).$$

将 $\tilde{\alpha}_3$ 标准化为 α_3，即令

$$\alpha_3 = \frac{1}{\sqrt{\tilde{\alpha}_3\tilde{\alpha}_3'}}\tilde{\alpha}_3 = \frac{1}{2}(0, 0, 2) = (0, 0, 1).$$

总起来，便得

$$\alpha_1 = \frac{1}{\sqrt{2}}\beta_1$$

$$\alpha_2 = \qquad \frac{1}{\sqrt{2}}\beta_2$$

$$\alpha_3 = \frac{1}{2}\left(-\frac{1}{\sqrt{2}}\alpha_1 + \frac{1}{\sqrt{2}}\alpha_2 + \beta_3\right)$$

$$= -\frac{1}{4}\beta_1 + \frac{1}{4}\beta_2 + \frac{1}{2}\beta_3$$

于是写成矩阵形式就是

$$T = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

即

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

使 $PA = T$ 为正交矩阵。

例 2 设 $\alpha = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$。求一个正交矩阵$T$，使 α' 是 T 的第一列。

解 可用两种办法来做。

方法一。任取三维向量 β, γ，使

$$A = (\alpha' \beta' \gamma')$$

为可逆阵。接下去按例 1 的方法求出T即可。因为这里的 α题设就是单位向量，所以它就是第一步求出的那个单位向量，因而也就是最后所求那个正交矩阵的第一列。

方法二。考虑以 $\alpha = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$为系数阵的齐次线性方程组

$$\left(\frac{1}{\sqrt{2}}\ 0\ -\frac{1}{\sqrt{2}}\right)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

求出它的一个基础解系。比如

$$\beta=(0,1,0),\gamma=(1,0,1).$$

于是 $\alpha\perp\beta,\alpha\perp\gamma$，而 $\beta\perp\gamma$.

以下把 β,γ 标准化. 取

$$\beta_1=\frac{1}{\sqrt{\beta\beta'}}\beta=\beta=(0,1,0)$$

$$\gamma_1=\frac{1}{\sqrt{\gamma\gamma'}}\gamma=\frac{1}{\sqrt{2}}(1,0,1)=\left(\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right).$$

这样

$$T=(\alpha'\beta'\gamma')=\begin{pmatrix}\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\\ 0 & 1 & 0\\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}}\end{pmatrix}$$

就是正交矩阵，其第一列恰为 α'.

练　习　四

1. 验证下列矩阵为正交矩阵

$$A=\begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix},\quad B=\begin{pmatrix}1 & & & \\ & -1 & & \\ & & \cos\alpha & \sin\alpha\\ & & -\sin\alpha & \cos\alpha\end{pmatrix},$$

$$C=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & \frac{1}{2}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2}\\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & -\frac{1}{2}\\ 0 & 0 & \frac{\sqrt{3}}{2} & \frac{1}{2}\end{pmatrix}.$$

$$D=\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&-1&&\\&&&&\ddots&\\&&&&&-1\end{pmatrix}.$$

2．设

$$A=\begin{pmatrix}a_1&&\\&\ddots&\\&&a_n\end{pmatrix}.$$

证明：A为正交矩阵的充分必要条件是 $a_i=\pm1$，$i=1,2,\cdots,n$。

3．设 A、P 都是正交矩阵．证明 $P^{-1}AP$也是正交矩阵．举例说明：A为正交矩阵，P不是正交矩阵，但 $P^{-1}AP$ 还是正交矩阵．

4．设A为对称阵，P为正交矩阵．证明 $P^{-1}AP$ 为对称阵．

5．设 $A=(a_{ij})$ 是行列式为1的三阶正交阵．证明：

$$a_{11}=a_{22}a_{33}-a_{23}a_{32}$$

$$a_{21}=a_{13}a_{32}-a_{12}a_{33}$$

$$a_{31}=a_{12}a_{23}-a_{13}a_{22}.$$

6．设 $A=(a_{ij})$ 为 n 阶正交阵．证明：方程组

$$A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

的解为

$$\begin{cases}x_1=a_{11}b_1+a_{21}b_2+\cdots+a_{n1}b_n\\x_2=a_{12}b_1+a_{22}b_2+\cdots+a_{n2}b_n\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\x_n=a_{1n}b_1+a_{2n}b_2+\cdots+a_{nn}b_n.\end{cases}$$

7．证明：正交阵的任一行（列）乘以－1后，结果仍为正交阵．

8．证明下列的三个条件中只要有两个成立，那么另一个也成立：

1）A是对称阵；

2）A是正交阵；

3）A是对合阵。

9．设A为实对称阵，S为反对称阵①，且 $AS=SA$及 $A-S$ 为可逆阵．证明：$(A+S)(A-S)^{-1}$ 为正交阵。

10．把下列向量组标准正交化：

1）$\alpha_1=(1,0,0)$，$\alpha_2=(1,2,0)$，$\alpha_3=(1,1,3)$，

2）$\beta_1=(1,1,0,0)$，$\beta_2=(1,0,1,0)$，$\beta_3=(-1,0,0,2)$，$\beta_4=(1,-1,-1,1)$。

11．设

$$A=\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ 0 & 1 & 2\end{pmatrix}$$

试求一三角形阵P，使 PA为正交阵。

12．设向量组 $\alpha_1=(2,1,0,1,-3)$，$\alpha_2=(1,1,-1,0,2)$．试求 $\alpha_3,\alpha_4,\alpha_5$，使 α_1,α_2与$\alpha_3,\alpha_4,\alpha_5$成正交。

13．求齐次线性方程组

$$\begin{cases}2x_1+x_2-x_3+x_4-3x_5=0\\ x_1+x_2-x_3+x_5=0\end{cases}$$

的解空间的一组标准正交基底。

14．设$e_1,e_2,\cdots,e_n$ 为 n维向量空间的标准正交基底．$T=(t_{ij})$ 为正交阵．而

$\varepsilon_1=t_{11}e_1+t_{21}e_2+\cdots+t_{n1}e_n$，$\varepsilon_2=t_{12}e_1+t_{22}e_2+\cdots+t_{n2}e_n$，$\cdots$，$\varepsilon_n=t_{1n}e_1+t_{2n}e_2+\cdots+t_{nn}e_n$。

证明 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 也是标准正交基底。

15．设 n维向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r\ (r<n)$ 是标准正交组．证明：可以找到 $n-r$个 n维向量 $\alpha_{r+1},\cdots,\alpha_n$，使 $\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n$ 为标准正交基底。

---

①反对称阵的定义：一个方阵A叫做反对称阵，如果 $A'=-A$。

§5　实对称阵在正交相合之下的标准形

本节确定实对称阵在正交相合之下的标准形．我们将证明：任一实对称阵A，存在正交阵P，使 $P'AP=P^{-1}AP$ 为对角形阵．

我们已经知道：方阵A相似于对角形阵的充分必要条件是A有完全的特征向量系．由上节的讨论又知道：A正交相合（相似）于对角形阵的充分必要条件是A有完全的特征向量系，并且它又是标准正交的．也就是A有完全的标准正交的特征向量系．

因此，我们就来证明：实对称阵必有完全标准正交的特征向量系．

命题1　**设A为实对称阵．那么A的特征根都是实数．**

证明　设λ为A的任一特征根．任取属于特征根λ的特征向量

$$\alpha=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},\quad a_i\text{ 不全为零．}$$

于是

$$A\alpha=\lambda\alpha.$$

两端分别取转置和共轭，便有

$$\overline{\alpha'}\,\overline{A'}=\overline{\lambda}\,\overline{\alpha'}.$$

因为A是实对称阵，所以$\overline{A'}=A$，故得

$$\overline{\alpha'}A=\overline{\lambda}\,\overline{\alpha'}.$$

用α从右侧去乘上式，则有

$$\overline{\alpha'}A\alpha=\overline{\lambda}\,\overline{\alpha'}\alpha$$

即　$$\overline{\alpha'}(\lambda\alpha)=\overline{\lambda}\,\overline{\alpha'}\alpha\text{或}\lambda\overline{\alpha'}\alpha=\overline{\lambda}\,\overline{\alpha'}\alpha.$$

由于 $\alpha\neq0$，所以$\overline{\alpha'}\alpha\neq0$，于是便得

$$\lambda=\overline{\lambda}$$，即 λ为实数．证完．

命题2　**设A为实对称阵，那么A的属于不同特征根的特征向**

量是正交的。

证明　设 λ_1, λ_2 是A的两个不同的特征根，α, β分别为属于 λ_1, λ_2 的特征向量：

$$A\alpha = \lambda_1\alpha, \quad A\beta = \lambda_2\beta$$

于是

$$\alpha'A' = \lambda_1\alpha'.$$

用 β 从右侧去乘上式两端，则有

$$\alpha'A'\beta = \lambda_1\alpha'\beta \text{ 或 } \alpha'A\beta = \lambda_1\alpha'\beta.$$

把 $A\beta = \lambda_2\beta$ 代入上式，便有

$$\lambda_2\alpha'\beta = \lambda_1\alpha'\beta$$

因为 $\lambda_1 \neq \lambda_2$，所以可得

$\alpha'\beta = 0$，即 α 与 β 正交，证完。

定理 1　设A为实对称阵，则A必有完全的特征向量系。

证明　根据特征向量系的定义，只须证明：对A的任一特征根 λ，如果 λ 的重数等于 $S_A(\lambda)$的维数即可。

设λ 的重数为s，$S_A(\lambda)$ 的维数为r。在 $S_A(\lambda)$ 中任取一标准正交基底：

$$u_1, u_2, \cdots, u_r.$$

把它补充成整个空间 V_n(n 维向量空间）的一个标准正交基底：

$$u_1, u_2, \cdots, u_r, u_{r+1}\cdots, u_n.$$

于是可以写出以下 n个等式：

$$\begin{aligned}
Au_1 &= \lambda u_1 \\
Au_2 &= \quad\quad \lambda u_2 \\
&\vdots \quad\quad\quad \ddots \\
Au_r &= \quad\quad\quad\quad\quad \lambda u_r \\
Au_{r+1} &= a_{1r+1}u_1 + \cdots + a_{rr+1}u_r + a_{r+1r+1}u_{r+1} + \cdots + a_{nr+1}u_n \\
&\vdots \\
Au_n &= a_{1n}u_1 + \cdots + a_{rn}u_r + a_{r+1n}u_{r+1} + \cdots + a_{nn}u_n.
\end{aligned}$$

把这 n个等式拼写成一个矩阵的等式，就是

$$A(u_1u_2\cdots u_ru_{r+1}\cdots u_n)$$

$$=(u_1u_2\cdots u_ru_{r+1}\cdots u_n)\left(\begin{array}{ccc|ccc}\lambda & & & a_{1r+1} & \cdots & a_{1n}\\ & \lambda & & \vdots & & \vdots \\ & & \ddots\ \lambda & a_{rr+1} & \cdots & a_{rn}\\ \hline & & & a_{r+1r+1} & \cdots & a_{r+1n}\\ & 0 & & \vdots & & \\ & & & a_{nr+1} & \cdots & a_{nn}\end{array}\right).$$

由 $P=(u_1u_2\cdots u_ru_{r+1}\cdots u_n)$ 是正交阵，可知 $P^{-1}AP$ 是对称阵. 因此就有

$$A(u_1u_2\cdots u_ru_{r+1}\cdots u_n)$$

$$=(u_1u_2\cdots u_ru_{r+1}\cdots u_n)\left(\begin{array}{c|c}\begin{matrix}\lambda & & \\ & \lambda & \\ & & \ddots\ \lambda\end{matrix} & 0\\ \hline 0 & A_1\end{array}\right),$$

此式中的 A_1 是 $n-r$阶实对称阵.

我们只要证明 A_1 不再以 λ 为特征根，便可断定 $r=s$.

用反证法. 假设 λ 还是 A_1 的特征根，那么必有 $(n-r)\times 1$的非零矩阵 β，使

$$A_1\beta=\lambda\beta.$$

类似地可知，存在 $n-r$ 阶正交阵 Q_1，使

$$A_1Q_1=Q_1\left(\begin{array}{c|c}\lambda & 0\\ \hline 0 & A_2\end{array}\right).$$

令

$$Q=\begin{pmatrix}\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix}\Big\}r\text{个} & \\ & Q_1\end{pmatrix}.$$

于是

$$\left(\begin{array}{c|c}\begin{matrix}\lambda & & \\ & \ddots & \\ & & \lambda\end{matrix}\Big\}r\text{个} & 0\\ \hline 0 & A_1\end{array}\right)Q=Q\left(\begin{array}{c|c}\begin{matrix}\lambda & & \\ & \ddots & \\ & & \lambda \\ & & & \lambda\end{matrix}\Big\}r+1\text{个} & 0\\ \hline 0 & A_2\end{array}\right).$$

从而

$$A(u_1\cdots u_r u_{r+1}\cdots u_n)Q$$

$$=(u_1\cdots u_r u_{r+1}\cdots u_n)Q\left(\begin{array}{c|c}\begin{matrix}\lambda & & \\ & \ddots & \\ & & \lambda \\ & & & \lambda\end{matrix}\ \overbrace{}^{r+1\text{个}} & 0 \\ \hline 0 & A_2\end{array}\right).$$

其中 $(u_1\cdots u_r u_{r+1}\cdots u_n)Q=PQ$ 仍是正交阵，而前 $r+1$ 个列向量都是 A 的属于同一特征根 λ 的特征向量。这与 $S_A(\lambda)$ 是 r 维的相矛盾，由此说明上边的 A_1 不能再有特征根 λ。故得 r 就是 λ 的重数，即 $S_A(\lambda)$ 的维数 r 等于 λ 的重数 s。证完。

推论　一个实对称阵 A，总存在可逆阵 P，使

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 A 的全部特征根。

定理 2　设 A 为实对称阵，那么 A 必有完全的标准正交的特征向量系。

证明　设 $\lambda_1,\lambda_2,\cdots,\lambda_t$ 为 A 的一切不同的特征根，且令 $S_A(\lambda_1)$，$S_A(\lambda_2),\cdots,S_A(\lambda_t)$ 为 A 的特征空间。由定理 1 知，

$$\dim S_A(\lambda_1)+\dim S_A(\lambda_2)+\cdots+\dim S_A(\lambda_t)=n\ (A\text{ 的阶数}).$$

在各特征空间中任取标准正交基底：

$$u_{11}\cdots u_{1l_1};\ u_{21},\cdots,u_{2l_2};\ \cdots,\ u_{t1},\cdots,u_{tl_t}$$

由命题 2 知，这就是 A 的一个完全标准正交的特征向量系。证完。

于是，当设

$$P=(u_{11}\cdots u_{1l_1}u_{21}\cdots u_{2l_2}\cdots u_{t1}\cdots u_{tl_t})$$

时，便有

$$P^{-1}AP=\begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & & & & & & \ddots & \\ & & & & & & & \lambda_t \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_t \end{pmatrix} \begin{matrix} \left.\right\} l_1\text{个} \\ \left.\right\} l_2\text{个} \\ \left.\right\} l_t\text{个} \end{matrix}$$

这个对角形阵叫做实对称阵A的标准形。

推论 1　一个实对称阵A，总存在正交阵P，使

$$P'AP=P^{-1}AP=\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

其中 $\lambda_1,\cdots,\lambda_n$ 为A的全部特征根。

推论 2　实对称阵A是正（负）定的充分必要条件为其特征根都是正（负）数；实对称阵A是半正（负）定的充分必要条件为其特征根都是非负（正）数。

推论 3　对于任一实二次型

$$f(x_1,\cdots,x_n)=(x_1\cdots x_n)A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

都有适当的变数变换

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}=P\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

使 $f(x_1,\cdots,x_n)$化为平方和：

$$g(y_1,\cdots,y_n)=\lambda_1y_1^2+\cdots+\lambda_ny_n^2,$$

其中P为正交阵，$\lambda_1,\cdots,\lambda_n$ 为A的全部特征根。

这种化二次型为平方和的方法叫做正交变换法。

例 1　设三阶实对称阵

$$A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix},$$

求正交矩阵 T，使 $T'AT=T^{-1}AT$ 为对角形阵。

解　这个对称阵 A 在本章§3已求出了它的完全特征向量系：

$$\alpha_1=(1,0,-1),\quad \alpha_2=(0,1,-1),\quad \alpha_3=(1,1,1).$$

其中 α_1,α_2 是属于特征根 -1 的特征空间 $S_A(-1)$ 的一个基底；α_3 是属于特征根 5 的特征空间 $S_A(5)$ 的一个基底。并且有

$$(\alpha_1'\alpha_2'\alpha_3')^{-1}A(\alpha_1'\alpha_2'\alpha_3')=\begin{pmatrix}-1&&\\&-1&\\&&5\end{pmatrix}.$$

这里演化矩阵 $P=(\alpha_1'\alpha_2'\alpha_3')$ 还不是正交矩阵，即特征向量系 $\alpha_1,\alpha_2,\alpha_3$ 是完全的，但还不是标准正交的。为了得到 A 的完全标准正交特征向量系（当然也就是使演化矩阵为正交阵），只须把 $\alpha_1,\alpha_2,\alpha_3$ 分别标准正交化即可。

先标准正交化 α_1,α_2。

令

$$\beta_1=\frac{1}{\sqrt{\alpha_1\alpha_1'}}\alpha_1=\frac{1}{\sqrt{2}}(1,0,-1)=\left(\frac{1}{\sqrt{2}},0,\frac{-1}{\sqrt{2}}\right)$$

再令

$$\beta_2^*=c_{21}\beta_1+\alpha_2,\ \text{使}\beta_2^*\perp\beta_1,\ \text{即}$$

$$\beta_2^*\beta_1'=(c_{21}\beta_1+\alpha_2)\beta_1'=c_{21}+\alpha_2\beta_1'=0.$$

于是

$$c_{21}=-\alpha_2\beta_1'=-(0,1,-1)\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{-1}{\sqrt{2}}\end{pmatrix}=\frac{-1}{\sqrt{2}},\ \text{故有}$$

$$\beta_2^*=\frac{-1}{\sqrt{2}}\beta_1+\alpha_2=\frac{-1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}},0,\frac{-1}{\sqrt{2}}\right)+(0,1,-1)=\left(\frac{-1}{2},1,-\frac{1}{2}\right)$$

取 $\beta_2=\frac{1}{\sqrt{\beta_2\beta_2'}}\beta_2^*=\frac{\sqrt{2}}{\sqrt{3}}\left(\frac{-1}{2},1,\frac{-1}{2}\right)=\left(\frac{-1}{\sqrt{6}},\frac{\sqrt{2}}{\sqrt{3}},\frac{-1}{\sqrt{6}}\right)$,

则得 β_1,β_2 为 $S_A(-1)$的一个标准正交基底。

再标准化 α_3。

令 $\beta_3=\frac{1}{\sqrt{\alpha_3\alpha_3'}}\alpha_3=\frac{1}{\sqrt{3}}(1,1,1)=\left(\frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}}\right)$,

即 β_3 为$S_A(5)$的标准正交基底。

总起来便得A的一个完全标准正交特征向量系：

$$\beta_1=\left(\frac{1}{\sqrt{2}},\ 0,\ \frac{-1}{\sqrt{2}}\right)$$

$$\beta_2=\left(\frac{-1}{\sqrt{6}},\ \frac{\sqrt{2}}{\sqrt{3}},\ \frac{-1}{\sqrt{6}}\right)$$

$$\beta_3=\left(\frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}},\ \frac{1}{\sqrt{3}}\right).$$

从而，若令

$$T=(\beta_1'\beta_2'\beta_3'),$$

则

$$T^{-1}AT=T'AT=\begin{pmatrix}-1 & & \\ & -1 & \\ & & 5\end{pmatrix}.$$

例 2 设实二次型 $f(x_1,x_2,x_3)=x_1^2+4x_1x_2+4x_1x_3+x_2^2+4x_2x_3+x_3^2$，试用正交变换法化为平方和。

解 二次型 $f(x_1,x_2,x_3)$的表示矩阵为

$$A=\begin{pmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{pmatrix}.$$

由上例可知，存在正交矩阵

$$T=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}}\\ \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}}\end{pmatrix}.$$

使

$$T'AT=T^{-1}AT=\begin{pmatrix}-1 & & \\ & -1 & \\ & & 5\end{pmatrix}.$$

于是经可逆变换

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=T\begin{pmatrix}y_1\\y_2\\y_2\end{pmatrix}$$

使二次型

$$f(x_1,x_2,x_3)=g(y_1,y_2,y_3)=-y_1^2-y_2^2+5y_3^2$$

练　习　五

1．求下列对称阵的完全标准正交的特征向量系

$$(1)\begin{pmatrix}2 & -2 & 0\\ -2 & 1 & -2\\ 0 & -2 & 0\end{pmatrix}\qquad (2)\begin{pmatrix}-1 & -3 & 3 & -3\\ -3 & -1 & -3 & 3\\ 3 & -3 & -1 & -3\\ -3 & 3 & -3 & -1\end{pmatrix}$$

2．求正交阵 P、Q，使 $P^{-1}AP$，$Q^{-1}BQ$ 为对角形，并指出A，B是否为正定的，其中

$$A=\begin{pmatrix}2 & 2 & -2\\ 2 & 5 & -4\\ -2 & -4 & 5\end{pmatrix},\quad B=\begin{pmatrix}0 & 0 & 4 & 1\\ 0 & 0 & 1 & 4\\ 4 & 1 & 0 & 0\\ 1 & 4 & 0 & 0\end{pmatrix}.$$

3．用正交变换化下列二次型为平方和，并指出变数变换的正交矩阵。

1）$x_1^2+2x_2^2+3x_3^2-4x_1x_2-4x_1x_3$

2）$x_1^2+x_2^2+x_3^2+x_4^2-2x_1x_2+6x_1x_3-4x_1x_4-4x_2x_3+6x_2x_4-2x_3x_4$.

4．设 A，B 都是实对称阵．证明，存在正交矩阵T，使$T^{-1}AT=B$的充分必要条件是A与B有相同的特征根．

5．证明：幂零阵又是实对称阵必是零阵．

6．证明：对称的正交矩阵A的特征根必为 1 或 -1．

7．设 λ、μ 为对称阵A的不同特征根，证明：$(\lambda I-A)X=0$ 与 $(\mu I-A)Y=0$ 的非零解恒正交．

8．设 A，B 均为实对称阵，且 B 为正定矩阵．证明，存在实可逆阵T，使

$$T'AT \text{ 与 } T'BT$$

同时为对角形．

9．设 A，B 均为正定矩阵．证明，A，B 的特征根均为正的．

§6 正交矩阵在正交相似之下的标准形

设 A、T 都是正交矩阵，那么 $T^{-1}AT=T'AT$ 也是正交矩阵，并且正交相似满足反身性，对称性和传递性三个条件．这样，全体正交矩阵按正交相似（相合）的关系也做成了分类．换句话说，以正交矩阵为演化矩阵，正交矩阵本身也存在着化简问题：对给定的正交矩阵A，如何选取正交矩阵T，可使 $T^{-1}AT=T'AT$得到怎样的简单形式？

本节就是讨论正交矩阵在正交相似之下的标准形问题．所用的方法基本上是求特征向量的方法．

命题 1　设A为 n 阶正交阵，λ_0 为A的任一实特征根，那么 $\lambda_0=\pm1$，且存在正交阵P，使

$$P^{-1}AP=P'AP=\begin{pmatrix}\lambda_0 & 0\ \cdots\ 0\\ 0 & \\ \vdots & A_1\\ 0 & \end{pmatrix},$$

其中 A_1 为 $n-1$ 阶正交阵．

证明　因为正交阵A的特征根的模等于1，所以A的实特征根 $\lambda_0=1$ 或 -1．

任取属于 λ_0 的一个特征向量

$$\bar{a}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\qquad a_i\text{全为实数(这是可以做到的)}$$

则有

$$A\bar{a}=\lambda_0\bar{a}.$$

取 $\alpha=\dfrac{1}{\sqrt{\bar{a}\bar{a}'}}\bar{a}$，则 $|\alpha|=1$且有 $A\alpha=\lambda_0\alpha$。令 P 为以 α 为第一列的一个正交阵，即 $P=(\alpha\ \beta_2\cdots\beta_n)$。由于

$$(P'A)\alpha=P'(A\alpha)=P'(\lambda_0\alpha)=\lambda_0(P'\alpha)=\lambda_0\begin{pmatrix}\alpha'\\\beta'_2\\\vdots\\\beta'_n\end{pmatrix}\alpha=$$

$$=\lambda_0\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}=\begin{pmatrix}\lambda_0\\0\\\vdots\\0\end{pmatrix}.$$

所以

$$P^{-1}AP=P'AP=P'A(\alpha\beta_2\cdots\beta_n)=\begin{pmatrix}\lambda_0&b_2&\cdots&b_n\\0&&&\\\vdots&&A_1&\\0&&&\end{pmatrix}.$$

因为 $P^{-1}AP=P'AP$ 也是正交矩阵，$\lambda_0=\pm1$，而正交矩阵第一行元素平方之和等于1，故推知：

$$b_2=0,\cdots,b_n=0.$$

从而

$$P^{-1}AP=P'AP=\begin{pmatrix}\lambda_0&0&\cdots&0\\0&&&\\\vdots&&A_1&\\0&&&\end{pmatrix},$$

并且 A_1 是 $n-1$ 阶正交阵，证完。

命题2　设A为正交阵，$\lambda_0=u+vi$是A的任一虚特征根（即$v\neq0$），那么存在正交阵P，使

$$P^{-1}AP = P'AP = \begin{pmatrix} A_2 & & 0 & \\ & 0 & \cos\theta & -\sin\theta \\ & & \sin\theta & \cos\theta \end{pmatrix}.$$

其中 A_2 是$n-2$阶正交阵.

证明　任取属于特征根 λ_0 的特征向量

$$\gamma = \begin{pmatrix} a_1 + b_1 i \\ a_2 + b_2 i \\ \vdots \\ a_n + b_n i \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} i$$

令

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$ 这是两个实 $n\times 1$矩阵.

于是，由 $A\gamma = \lambda_0\gamma$，可得

$$A(\alpha + \beta i) = (u + vi)(\alpha + \beta i),$$

即有 $$A\alpha + A\beta i = (u\alpha - v\beta) + (v\alpha + u\beta)i,$$

从而

$$A\alpha = u\alpha - v\beta$$

$$A\beta = v\alpha + u\beta.$$

把上边两个式子拼写成一个矩阵等式

$$A(\alpha\beta) = (\alpha\beta)\begin{pmatrix} u & v \\ -v & u \end{pmatrix}.$$

现在我们指出：α,β 看做两个 n维实向量是线性无关的.

首先容易知道，α,β 都不是零向量. 比如，若 $\alpha=0$，由 $A\alpha = u\alpha - v\beta$ 可得 $v\beta=0$，但 $v\neq 0$，所以就有 $\beta=0$ 这与 $\gamma=\alpha+\beta i$ 是特征向量矛盾，故 $\alpha\neq 0$.

同理也有 $\beta\neq 0$.

假设 α,β 线性相关，可令 $\beta=k\alpha$. 于是

$$A\alpha = u\alpha - v\beta = u\alpha - v(k\alpha) = (u-kv)\alpha$$

$$A\beta = v\alpha + u\beta = v\alpha + u(k\alpha) = (v+ku)\alpha.$$

而 $A\beta = kA\alpha$. 由此可得

$$k(u-kv)\alpha=(v+ku)\alpha.$$

因$\alpha\neq 0$，所以 $ku-k^2v=v+ku$。从而得 $(k^2+1)v=0$。又因 $v\neq 0$，$k^2+1\neq 0$，这样又是一个矛盾。因之 α,β 不能是线性相关的，故必线性无关。

既然 α,β 是线性无关的，可取二阶上三角形阵 B，使

$$(\alpha\beta)B=(\bar{a}\beta^*)\text{ 中的}\bar{a},\beta^*$$

是标准正交的，这样任取 n 阶正交矩阵 P，使其最后两列为$\bar{a},\beta^*$，即$P=(\alpha_1\cdots\alpha_{n-2}\bar{a}\beta^*)$。则有

$$P^{-1}AP=P'AP=\begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha'_{n-2}\\ \bar{a}'\\ \beta^*\end{pmatrix}A\ (\alpha_1\cdots\alpha_{n-2}\bar{a}\beta^*)$$

$$=\begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha'_{n-2}\\ \hline \bar{a}'\\ \beta^*\end{pmatrix}(A(\alpha_1\cdots\alpha_{n-2})A(\bar{a}\beta^*))$$

$$=\begin{pmatrix}\begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha'_{n-2}\end{pmatrix}A(\alpha_1\cdots\alpha_{n-2}) & \begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha'_{n-2}\end{pmatrix}A(\bar{a}\beta^*)\\ \begin{pmatrix}\bar{a}'\\ \beta^*\end{pmatrix}A\ (\alpha_1\cdots\alpha_{n-2}) & \begin{pmatrix}\bar{a}'\\ \beta^*\end{pmatrix}A(\bar{a}\beta^*)\end{pmatrix}=\begin{pmatrix}A_2 & \begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha'_{n-2}\end{pmatrix}A\ (\alpha\beta)\ B\\ A_3 & \begin{pmatrix}\bar{a}'\\ \beta^*\end{pmatrix}A\ (\alpha\beta)\ B\end{pmatrix}$$

$$=\begin{pmatrix}A_2 & \begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha'_{n-2}\end{pmatrix}(\alpha\beta)\begin{pmatrix}u & v\\ -v & u\end{pmatrix}B\\ A_3 & \begin{pmatrix}\bar{a}'\\ \beta^*\end{pmatrix}(\alpha\beta)\begin{pmatrix}u & v\\ -v & u\end{pmatrix}B\end{pmatrix}=\begin{pmatrix}A_2 & \begin{pmatrix}\alpha_1'\\ \vdots\\ \alpha'_{n-2}\end{pmatrix}(\bar{a}\beta^*)B^{-1}\begin{pmatrix}u & v\\ -v & u\end{pmatrix}B\\ A_3 & \begin{pmatrix}\bar{a}'\\ \beta^*\end{pmatrix}(\bar{a}\beta^*)B^{-1}\begin{pmatrix}u & v\\ -v & u\end{pmatrix}B\end{pmatrix}$$

$$=\begin{pmatrix}A_2 & \begin{pmatrix}0 & 0\\ \vdots & \vdots\\ 0 & 0\end{pmatrix}C\\ A_3 & C\end{pmatrix}=\begin{pmatrix}A_2 & 0\\ A_3 & C\end{pmatrix},\quad \text{其中 } C=B^{-1}\begin{pmatrix}u & v\\ -v & u\end{pmatrix}B.$$

因 $P^{-1}AP$ 为正交阵，C是可逆的：$|C|=\left|B^{-1}\begin{pmatrix}u & v\\ -v & u\end{pmatrix}B\right|$

$=\begin{vmatrix} u & v \\ -v & u \end{vmatrix}=u^2+v^2=1$。由此便可推知 $2\times(n-2)$ 矩阵 $A_3=0$。

事实上，令

$$A_3=\begin{pmatrix} c_1 & c_2 & \cdots & c_{n-2} \\ d_1 & d_2 & \cdots & d_{n-2} \end{pmatrix}.$$

由 $P^{-1}AP=P'AP$ 的列的正交性，可得：

$$(c_1\ d_1)C=0,\quad (c_2\ d_2)C=0,\cdots,\quad (c_{n-2}\ d_{n-2})C=0.$$

以上这 $n-2$ 个矩阵等式表明：$(c_1\ d_1)$，$(c_2 d_2)$，…，$(c_{n-2} d_{n-2})$ 都是齐次线性方程组

$$(x_1 x_2)C=0$$

的解．因 $|C|\neq 0$，所以这个齐次线性方程组只有零解，故

$$c_1=0,\ d_1=0;\ c_2=0,\ d_2=0;\cdots;\ c_{n-2}=0,\ d_{n-2}=0,$$

即

$$A_3=\begin{pmatrix} c_1 & c_2 & \cdots & c_{n-2} \\ d_1 & d_2 & \cdots & d_{n-2} \end{pmatrix}=0.$$

这样，就有

$$P^{-1}AP=P'AP=\begin{pmatrix} A_2 & 0 \\ 0 & C \end{pmatrix}$$

为正交阵，从而其中的 A_2，C分别是 $n-2$ 阶正交阵与2阶正交阵。

又因$C=B^{-1}\begin{pmatrix} u & v \\ -v & u \end{pmatrix}B$，所以 $|C|=u^2+v^2=1$。由本章§4例 3 知，

$$C=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

即有

$$P^{-1}AP=P'AP=\begin{pmatrix} A_2 & & 0 \\ 0 & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}.$$ 证完。

注　命题1的意思是说：对n阶正交阵A来说，通过它的每一个实特征根λ_0，都可以把A简化一步，使

$$A \sim \begin{pmatrix} \lambda_0 & \\ & A_1 \end{pmatrix},$$

其中A_1是$n-1$阶正交阵。这样，化简A的问题就归结为去化简阶数已减少1的正交阵A_1。如此推理下去，就可通过实特征根，使n阶正交阵化简到这样的地步：

$$A \sim \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & -1 & \\ & & & & & & A_0 \end{pmatrix},$$

其中A_0是正交阵，它没有实特征根。

我们看到，利用正交矩阵的实特征根化简正交阵的基本思路和具体做法与化简实对称阵的情形完全一致。

命题2的意思是说：对n阶正交阵A来说，通过它的每一个虚特征根$\lambda_0=u+vi\,(v\neq 0)$都可以把$A$简化一步，使

$$A \sim \begin{pmatrix} A_2 & 0 & \\ 0 & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix},$$

其中A_2为$n-2$阶正交阵。这样化简A的问题就归结为去化简阶数已减少2的A_2，如此推理下去，就可通过虚特征根，使n阶正交阵A化简到这样的地步：

$$A \sim \begin{pmatrix} A_0 & & & & \\ & \cos\theta_1 & -\sin\theta_1 & & \\ & \sin\theta_1 & \cos\theta_1 & & \\ & & & \ddots & & \\ & & & & \cos\theta_r & -\sin\theta_r \\ & & & & \sin\theta_r & \cos\theta_r \end{pmatrix},$$

其中A_0是正交阵，它没有虚特征根。

再有，利用A的虚特征根$\lambda_0=u+vi\,(v\neq 0)$化简得

$$A\sim\begin{pmatrix}A_2 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta\end{pmatrix}$$

时，其中

$$C=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

满足

$$C\sim\begin{pmatrix}u & v\\ -v & u\end{pmatrix}.$$

所以C与$\begin{pmatrix}u & v\\ -v & u\end{pmatrix}$有相同的特征根。而

$$\begin{pmatrix}u & v\\ -v & u\end{pmatrix}\text{的特征根为 } u\pm vi,$$

$$C=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\text{的特征根为 } \cos\theta\pm i\sin\theta.$$

因此，$\cos\theta\mp i\sin\theta$ 分别是 $u\pm vi$ 的三角形式，故二阶正交阵C由已给的虚特征根 $\lambda_0=u+vi$ 所决定。

有了命题1和命题2及上述说明容易证明以下的

定理 **设A为正交阵，那么A正交相似于形式如下的矩阵，即存在正交阵T，使**

$$T^{-1}AT=\begin{pmatrix}1 &&&&&&&&&&\\ &\ddots&&&&&&&&&\\ &&1&&&&&&&&\\ &&&-1&&&&&&&\\ &&&&\ddots&&&&&&\\ &&&&&-1&&&&&\\ &&&&&&\cos\theta_1 & -\sin\theta_1&&&\\ &&&&&&\sin\theta_1 & \cos\theta_1&&&\\ &&&&&&&&\ddots&&\\ &&&&&&&&&\cos\theta_r & -\sin\theta_r\\ &&&&&&&&&\sin\theta_r & \cos\theta_r\end{pmatrix}, \quad (1)$$

并且可取 $\det T=1$.

证明 对A的阶数n用数学归纳法。当 $n=1$ 时，定理自然成立。假设对阶数小于n的正交阵定理成立，去证阶数为n的正交阵定理成立。

任取A的一个特征根 λ_0，如果λ_0是实数，由命题1，A正交相似于矩阵

$$\begin{pmatrix} \lambda_0 & \\ & A_1 \end{pmatrix}.$$

其中 A_1 是 $n-1$ 阶正交阵. 由归纳假设，有 A_1 正交相似于形如

$$\begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & \cos\theta_r & -\sin\theta_r \\ & & & & & & & & & \sin\theta_r & \cos\theta_r \end{pmatrix}$$

的矩阵，从而$\begin{pmatrix} \lambda_0 & \\ & A_1 \end{pmatrix}$正交相似于形如

$$\begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & \cos\theta_r & -\sin\theta_r \\ & & & & & & & & & \sin\theta_r & \cos\theta_r \end{pmatrix}$$

的矩阵. 故A正交相似于形式（1）的矩阵.

如果 λ_0 是虚数，由命题2，A正交相似于

$$\begin{pmatrix} A_2 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}$$

其中 A_2 为 $n-2$ 阶正交阵. 由归纳假设，有 A_2 正交相似于形如

$$\begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \begin{matrix} \cos\theta_{r-1} & -\sin\theta_{r-1} \\ \sin\theta_{r-1} & \cos\theta_{r-1} \end{matrix} \end{pmatrix}$$

的矩阵，从而

$$\begin{pmatrix} A_2 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}$$正交相似于形如

$$\begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & -1 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & -1 & & & & & \\ & & & & & & \begin{matrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{matrix} & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & \begin{matrix} \cos\theta_{r-1} & -\sin\theta_{r-1} \\ \sin\theta_{r-1} & \cos\theta_{r-1} \end{matrix} & \\ & & & & & & & & & \begin{matrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{matrix} \end{pmatrix}$$

的矩阵．故A正交相似于形式（1）的矩阵．也就是存在正交阵T，使

$$T^{-1}AT = T'AT = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \begin{matrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{matrix} \end{pmatrix}.$$

若 $\det T=-1$，则 $T'AT$ 两边乘以换法矩阵 C_{ii+1}：$C'_{ii+1}T'ATC_{ii+1}=(TC_{ii+1})'A(TC_{ii+1})=T_1'AT_1$，这时 $\det T_1=1$。故可假定原来的 T 为 $\det T=1$。本定理证完。

例 1 试决定二阶正交阵一切可能的标准形。

解 首先，若二阶正交阵有实特征根，那么它的两个特征根必都是实的，于是它的标准形必为下列形式之一：

$$\begin{pmatrix}1 & \\ & 1\end{pmatrix},\ \begin{pmatrix}1 & \\ & -1\end{pmatrix},\ \begin{pmatrix}-1 & \\ & -1\end{pmatrix}.$$

其次，如果二阶正交阵的特征根是虚数，那么它的标准形为：

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\quad 0<\theta<2\pi,\text{但 } \theta\neq\pi.$$

总之，二阶正交阵的标准形一切可能形式为：

$$\begin{pmatrix}1 & \\ & 1\end{pmatrix},\ \begin{pmatrix}1 & \\ & -1\end{pmatrix},\ \begin{pmatrix}-1 & \\ & -1\end{pmatrix},\ \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\ \theta\neq 0,\pi.$$

当然，其中的$\begin{pmatrix}1 & \\ & 1\end{pmatrix}$与$\begin{pmatrix}-1 & \\ & -1\end{pmatrix}$也能表成

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$

的形式，这只要 $\theta=0$ 或 $\theta=\pi$ 即可，在这种考虑之下，二阶正交阵的标准形就是下列两种形式之一：

$$\begin{pmatrix}1 & \\ & -1\end{pmatrix},\ \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},\quad 0\leqslant\theta<2\pi.$$

例 2 设正交阵

$$A=\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{pmatrix},$$

试求A的标准形及演化矩阵。

解 先求A的标准形，经计算知，A的特征根为$1,i,-i$。由虚根 i 的三角形式为

$$\cos\frac{3\pi}{2}-i\sin\frac{3\pi}{2}.$$

所以A的标准形为

$$\begin{pmatrix} 1 & & \\ & \cos\frac{3\pi}{2} & -\sin\frac{3\pi}{2} \\ & \sin\frac{3\pi}{2} & \cos\frac{3\pi}{2} \end{pmatrix}.$$

再求演化矩阵。

先求A的属于1的特征向量，即解方程组

$$(1I_3 - A)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \text{ 即}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

显然 $\varepsilon = (1,0,0)$ 为一个非零解向量，故 ε 为A的属于1的特征向量，并且 ε 是单位向量.于是所求的演化矩阵T，以 ε 为第一列。

其次求A的属于 i 的特征向量，即解方程组

$$(iI_3 - A)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \text{ 即}\begin{pmatrix} i-1 & 0 & 0 \\ 0 & i & 1 \\ 0 & -1 & i \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

显然 $\gamma = (0, -1, i)$ 为一个非零解向量，故 γ 为A的属于 i 的特征向量。而

$$\gamma = (0, -1, 0) + (0, 0, 1)i.$$

于是，令 $\alpha = (0, -1, 0)$，$\beta = (0, 0, 1)$ 并且 α 与 β 为标准正交的，故所要求的演化矩阵的最后两列为α，β. 因此，演化矩阵为

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

练 习 六

1. 试决定四阶正交阵的一切可能的标准形。

2. 求正交阵

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & 0 & \frac{-\sqrt{2}}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ \frac{\sqrt{2}}{2} & 0 & 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$$

的标准形及演化矩阵。

3．证明：正交阵属于不同实特征根的特征向量是正交的。

§7 有理标准形 若当标准形

前边两节以特征根和特征向量为桥梁，解决了两类特殊矩阵——实对称阵和正交矩阵在相似分类之下的标准形问题。

本节给出对一般数域上的任意方阵普遍可行的求相似标准形的方法，以及给出对复数域上任意方阵普遍可行的求相似标准形的方法，这里只介绍方法，理论证明留在下节去讲。

我们已知：对任一 n 阶方阵 $A=(a_{ij})$ 都唯一确定一个特征矩阵

$$\lambda I-A=\begin{pmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{pmatrix}.$$

这个特征矩阵与A的主要区别就是它的元素中含有参数 λ. 这样，$\lambda I-A$中的元素 $\lambda-a_{11}$，$\lambda-a_{22}$，…，$\lambda-a_{nn}$ 等都是 λ 的一次多项式。一般来说，把一个以 λ 的多项式为元素的矩阵，统称为多项式矩阵，也简称为 “λ—矩阵”。比如

$$\begin{pmatrix} \lambda^4+1 & 2\lambda-1 & 0 \\ 3 & \lambda^2+\lambda+5 & \lambda \\ \lambda^2 & -1 & 8\lambda \end{pmatrix},\quad \begin{pmatrix} \lambda & \lambda^2-1 & 2 & -1 \\ 2\lambda+1 & 3\lambda^2+\lambda-1 & -\lambda & \lambda^3 \\ 0 & \lambda-1 & -1 & \lambda^4-1 \end{pmatrix}$$

都是λ—矩阵；A的特征矩阵就是一个由A唯一决定的非常特殊的λ—矩阵。显然，元素都是数的矩阵都可以看做 λ—矩阵。以下用 $A(\lambda)$，$B(\lambda)$，…表示 λ—矩阵。

对于一般的λ—矩阵，可以和数元矩阵一样做加、减、乘法运算；可以同样的规定行列式、子式和秩数等概念，并且相应的一些结论也都是成立的，这里和以下一般都不另做证明，需要时便直接引用就是了．再者，对矩阵起决定性作用的初等变换也可完全类似地对一般λ—矩阵给出

定义 1）**倍法变换** 用任意不等于零的常数d去乘λ—矩阵的某一行（列）的每个元素；

2）**消法变换** 把λ—矩阵某一行（列）的$k(\lambda)$倍加于另一行(列)上．

这里值得注意的是，在消法变换里所用的倍数可以是任意的多项式$k(\lambda)$，而在倍法变换里所用的倍数必须是非零常数d，决不允许用非零次的多项式．

和数元矩阵一样，用若干次消法变换和倍法变换可以交换λ—矩阵的两行（列）．把交换两行或两列的变换也叫做**换法变换**．因此，交换λ—阵的两行（列）时，相当于施行若干次消法变换和倍法变换，故今后可直接进行换法变换．

定理 任意λ—矩阵$A(\lambda)(\neq 0)$用初等变换都能化简成以下形式：

$$\begin{pmatrix} d_1(\lambda) & & & & & \\ & d_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & d_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix},$$

其中$d_i(\lambda)$都是首系数为1的多项式，并且依次能整除：
$d_1(\lambda)|d_2(\lambda)|\cdots|d_r(\lambda)$．

本定理的证明放在下一节．

显然，r就是$A(\lambda)$的秩．$d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$叫做$A(\lambda)$的**不变因子**．如果$A(\lambda)$是$A$的特征矩阵，即$A(\lambda)=\lambda I-A$，那么$A(\lambda)=\lambda I-A$的不变因子叫做$A$的**不变因子**。

有了以上的引述，我们就来分别介绍求相似标准形的两种方法。

（一）有理标准形及其求法。

设A为任意的n阶方阵。求A的有理标准形的步骤如下：

1）求特征矩阵 $\lambda I-A$ 的不变因子。按前述定理，对 $\lambda I-A$ 用初等变换一定能求出不变因子来。由于 $\lambda I-A$的秩等于n，所以$\lambda I-A$的不变因子共有n个：$d_1(\lambda)$，$d_2(\lambda),\cdots,d_n(\lambda)$。

2）求伴侣矩阵。对任一首系数为1的m次多项式

$$g(\lambda)=\lambda^m+b_1\lambda^{m-1}+b_2\lambda^{m-2}+\cdots+b_{m-1}\lambda+b_m \quad (m>0)$$

都与下列形式相当简单的m阶矩阵是互相唯一决定的：

$$N=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -b_m \\ 1 & 0 & 0 & \cdots & 0 & -b_{m-1} \\ 0 & 1 & 0 & \cdots & \vdots & -b_{m-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & -b_2 \\ 0 & 0 & 0 & \cdots & 1 & -b_1 \end{pmatrix}.$$

把矩阵N叫做 $g(\lambda)$ 的伴侣矩阵，而称 $g(\lambda)$ 为矩阵N的伴侣多项式。

这样，如果已具体求出A的不变因子为

$$1,\cdots,1,d_1(\lambda),\ d_2(\lambda),\ \cdots,\ d_s(\lambda).$$

其中 $d_1(\lambda)$，$d_2(\lambda),\cdots$，$d_s(\lambda)$都是非零次的不变因子：

$$d_1(\lambda)=\lambda^{n_1}+a_{11}\lambda^{n_1-1}+a_{12}\lambda^{n_1-2}+\cdots+a_{1\,n_1-1}\lambda+a_{1n_1}$$

$$d_2(\lambda)=\lambda^{n_2}+a_{21}\lambda^{n_2-1}+a_{22}\lambda^{n_2-2}+\cdots+a_{2n_2-1}\lambda+a_{2n_2}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$d_s(\lambda)=\lambda^{n_s}+a_{s1}\lambda^{n_s-1}+a_{s2}\lambda^{n_s-2}+\cdots+a_{sn_s-1}\lambda+a_{sn_s}.$$

于是这s个不变因子各有一个伴侣矩阵依次为

$$N_1=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_{1n_1} \\ 1 & 0 & 0 & \cdots & 0 & -a_{1n_1-1} \\ 0 & 1 & 0 & \cdots & \cdot & \cdot \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & 0 & -a_{12} \\ 0 & 0 & 0 & \cdots & 1 & -a_{11} \end{pmatrix}$$

$$N_2=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_{2n_2} \\ 1 & 0 & 0 & \cdots & 0 & -a_{2n_2-1} \\ 0 & 1 & 0 & \cdots & \cdot & \cdot \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{22} \\ 0 & 0 & 0 & \cdots & 1 & -a_{21} \end{pmatrix}$$

$$\vdots$$

$$N_s=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_{sn_s} \\ 1 & 0 & 0 & \cdots & 0 & -a_{sn_s-1} \\ 0 & 1 & 0 & \cdots & \cdot & \cdot \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{s2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{s1} \end{pmatrix}.$$

总之对任一 n 阶方阵A，通过它的非零次的不变因子都能求出一组伴侣矩阵。

3）组成有理型矩阵。设 n 阶方阵A的不变因子为

$$1,\cdots,1,d_1(\lambda),d_2(\lambda),\ \cdots,d_s(\lambda).$$

$N_1,N_2,\cdots,N_s$ 依次是它们的非零次的不变因子的伴侣矩阵．于是，我们自然可以组合而成一个形式相当简单的 n 阶方阵，如

$$B=\begin{pmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_s \end{pmatrix}.$$

这是一个分块对角形矩阵，它对角线上的各小块恰好依次是A的非零次的不变因子 $d_1(x),d_2(\lambda),\cdots,\ d_s(\lambda)$ 的伴侣矩阵。把阵B叫做**阵A的有理型矩阵**。

这样，对任一 n 阶方阵A，都能求出一个有理型矩阵．我们的结论是：A与其有理型矩阵B相似，即$A\sim B$．这个结论的正确性放在下一节去证明．把与A相似的有理型矩阵叫做**A的有理标准形**。

总括起来便得：对任一 n 阶方阵A都有有理标准形，并且可以按上述三个步骤具体地求出A的有理标准形。

下面通过例子，来熟悉上述方法的三个步骤。

例1　求三阶方阵

$$A=\begin{pmatrix}-1 & -2 & 6\\ -1 & 0 & 3\\ -1 & -1 & 4\end{pmatrix}$$

的有理标准形。

解　1）求A的不变因子．即用初等变换化简A的特征矩阵

$$\lambda I-A=\begin{pmatrix}\lambda+1 & 2 & -6\\ 1 & \lambda & -3\\ 1 & 1 & \lambda-4\end{pmatrix}.$$

具体化简过程如下：

$$\lambda I-A=\begin{pmatrix}\lambda+1 & 2 & -6\\ 1 & \lambda & -3\\ 1 & 1 & \lambda-4\end{pmatrix}\xrightarrow{P_{23}(-1),P_{13}(-\lambda-1)}$$

$$\begin{pmatrix}0 & -\lambda+1 & -\lambda^2+3\lambda-2\\ 0 & \lambda-1 & -\lambda+1\\ 1 & 1 & \lambda-4\end{pmatrix}\xrightarrow{C_{13}}\begin{pmatrix}1 & 1 & \lambda-4\\ 0 & \lambda-1 & -\lambda+1\\ 0 & -\lambda+1 & -\lambda^2+3\lambda-2\end{pmatrix}$$

$$\xrightarrow[P_{12}(-1),P_{13}(-\lambda+4)]{}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda-1 & -\lambda+1\\ 0 & -\lambda+1 & -\lambda^2+3\lambda-2\end{pmatrix}$$

$$\xrightarrow[P_{23}(1)]{}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda-1 & 0\\ 0 & -\lambda+1 & -\lambda^2+2\lambda-1\end{pmatrix}\xrightarrow{P_{32}(1)}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda-1 & 0\\ 0 & 0 & -\lambda^2+2\lambda-1\end{pmatrix}$$

$$\xrightarrow{D_3(-1)}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda-1 & 0\\ 0 & 0 & \lambda^2-2\lambda+1\end{pmatrix}.$$

到此已经是

$$d_1(\lambda)=1,\quad d_2(\lambda)=\lambda-1,\quad d_3(\lambda)=\lambda^2-2\lambda+1$$

并且，有

$$d_1(\lambda)\mid d_2(\lambda)\mid d_3(\lambda).$$

所以，$d_1(\lambda),d_2(\lambda),d_3(\lambda)$ 就是A的不变因子．

2）求伴侣矩阵．非零次的不变因子有两个：$d_2(\lambda)=\lambda-1$，$d_3=\lambda^2-2\lambda+1$．于是它们的伴侣矩阵依次是

$$N_1=(1),\ N_2=\begin{pmatrix}0 & -1\\ 1 & 2\end{pmatrix}.$$

3）组成有理型矩阵．把求得的两个伴侣矩阵组合起来得一有理型矩阵：

$$B=\begin{pmatrix}N_1 & \\ & N_2\end{pmatrix}=\begin{pmatrix}1 & & \\ & 0 & -1\\ & 1 & 2\end{pmatrix}.$$

这样，有

$$A\sim B=\begin{pmatrix}1 & & \\ & 0 & -1\\ & 1 & 2\end{pmatrix}$$

即A的有理标准形就是 B.

例 2　求三阶方阵

$$A=\begin{pmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{pmatrix}$$

的有理标准形．

解　1）求 A的不变因子．即用初等变换化简A 的特征矩阵$\lambda I-A$，具体过程如下：

$$\lambda I-A=\begin{pmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & -2\\ -2 & -2 & \lambda-1\end{pmatrix}\xrightarrow{C_{12},\,D_1\left(-\frac{1}{2}\right)}\begin{pmatrix}1 & \frac{1}{2}(1-\lambda) & 1\\ \lambda-1 & -2 & -2\\ -2 & -2 & \lambda-1\end{pmatrix}$$

$$\xrightarrow{P_{21}(-\lambda+1),\,P_{31}(2)}\begin{pmatrix}1 & \frac{1}{2}(1-\lambda) & 1\\ 0 & \frac{1}{2}(\lambda^2-2\lambda-3) & -\lambda-1\\ 0 & -\lambda-1 & \lambda+1\end{pmatrix}$$

$$\xrightarrow[P_{12}\left(-\frac{1}{2}(1-\lambda)\right),\ P_{13}(-1)]{}\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{2}(\lambda^2-2\lambda-3) & -\lambda-1\\ 0 & -\lambda-1 & \lambda+1\end{pmatrix}$$

$$\xrightarrow[C_{23},\,D_2(-1)]{}\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda+1 & \frac{1}{2}(\lambda^2-2\lambda-3)\\ 0 & -\lambda-1 & -\lambda-1\end{pmatrix}\xrightarrow{P_{32}(1)}$$

$$\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda+1 & \frac{1}{2}(\lambda^2-2\lambda-3)\\ 0 & 0 & \frac{1}{2}(\lambda^2-4\lambda-5)\end{pmatrix}\xrightarrow[P_{23}\left(-\frac{1}{2}(\lambda-3)\right)]{D_3(2)}$$

$$\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda+1 & 0\\ 0 & 0 & \lambda^2-4\lambda-5\end{pmatrix}.$$

到此已有 $d_1(\lambda)=1$，$d_2(\lambda)=\lambda+1$，$d_3(\lambda)=\lambda^2-4\lambda-5$，且 $d_1(\lambda)|d_2(\lambda)|d_3(\lambda)$。所以，$d_1(\lambda), d_2(\lambda), d_3(\lambda)$就是$A$的不变因子。

2）求伴侣矩阵。

$d_2(\lambda)=\lambda+1$的伴侣矩阵为 $N_1=(-1)$，

$d_3(\lambda)=\lambda^2-4\lambda-5$ 的伴侣矩阵为$N_2=\begin{pmatrix}0 & 5\\ 1 & 4\end{pmatrix}$。

3）组成有理型矩阵：

$$B=\begin{pmatrix}-1 & & \\ & 0 & 5\\ & 1 & 4\end{pmatrix},\quad A\sim B。$$

即B为A的有理标准形。

例 3 求三阶方阵

$$A=\begin{pmatrix}1 & 2 & 0\\ -1 & 1 & -1\\ 0 & -2 & 1\end{pmatrix}$$

的有理标准形。

解 1）求不变因子：

$$\lambda I-A=\begin{pmatrix}\lambda-1 & -2 & 0\\ 1 & \lambda-1 & 1\\ 0 & 2 & \lambda-1\end{pmatrix}\xrightarrow{D_1(-1),C_{12}}\begin{pmatrix}1 & \lambda-1 & 1\\ 1-\lambda & 2 & 0\\ 0 & 2 & \lambda-1\end{pmatrix}$$

$$\xrightarrow{P_{21}(-1+\lambda)}\begin{pmatrix}1 & \lambda-1 & 1\\ 0 & \lambda^2-2\lambda+3 & \lambda-1\\ 0 & 2 & \lambda-1\end{pmatrix}\xrightarrow[P_{12}(-\lambda+1),\ P_{13}(-1)]{}$$

$$\begin{pmatrix}1 & 0 & 0\\ 0 & \lambda^2-2\lambda+3 & \lambda-1\\ 0 & 2 & \lambda-1\end{pmatrix}\xrightarrow{C_{23}}\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & \lambda-1\\ 0 & \lambda^2-2\lambda+3 & \lambda-1\end{pmatrix}\xrightarrow{D_2\left(\frac{1}{2}\right)}$$

$$\begin{pmatrix}1 & 0 & 0\\0 & 1 & \frac{1}{2}(\lambda-1)\\0 & \lambda^2-2\lambda+3 & \lambda-1\end{pmatrix}\xrightarrow[P_{23}\left(-\frac{1}{2}(\lambda-1)\right)]{}$$

$$\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & \lambda^2-2\lambda+3 & -\frac{1}{2}(\lambda-1)^3\end{pmatrix}\xrightarrow{P_{32}(-\lambda^2+2\lambda-3)}$$

$$\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & (\lambda-1)^3\end{pmatrix}.$$

于是不变因子为 $d_1(\lambda)=1$, $d_2(\lambda)=1$, $d_3(\lambda)=(\lambda-1)^3=\lambda^3-3\lambda^2+3\lambda-1$.

2）求伴侣矩阵. $d_3(\lambda)$ 的伴侣矩阵为

$$N_1=\begin{pmatrix}0 & 0 & 1\\1 & 0 & -3\\0 & 1 & 3\end{pmatrix}.$$

3）组成有理型矩阵:

$$B=N_1=\begin{pmatrix}0 & 0 & 1\\1 & 0 & -3\\0 & 1 & 3\end{pmatrix},\quad A\sim B.$$

即 B 为 A 的有理标准形.

（二）若当标准形及其求法.

设 A 为复数矩阵. 换句话说, 我们在复数域上考虑问题. 与求 A 的有理标准形一样, 这里也是三个步骤.

1）求初等因子. 先给出初等因子的定义:

对任意一个首系数为 1 的多项式 $f(\lambda)$, 在复数域上写出它的标准分解式:

$$f(\lambda)=(\lambda-c_1)^{l_1}(\lambda-c_2)^{l_2}\cdots(\lambda-c_r)^{l_r},$$

此处 $c_1,c_2,\cdots,c_r$ 是 r 个不同的复数, $l_1,l_2,\cdots,l_r$ 是非负整数. 我们把每一指数 l_i 大于零的幂 $(\lambda-c_i)^{l_i}$ 都叫做 $f(\lambda)$ 的初等因子.

这样, 若已求出 $A(\lambda)$ 的不变因子 $1,\cdots,1,d_1(\lambda),d_2(\lambda),\cdots,$

$d_r(\lambda)$，其中$d_1(\lambda),d_2(\lambda),\cdots,d_s(\lambda)$都是非零次的。于是对每一个$d_i(\lambda)$都可求出它的初等因子如下：

$$d_1(\lambda)=(\lambda-c_1)^{l_{11}}(\lambda-c_2)^{l_{12}}\cdots(\lambda-c_r)^{l_{1r}}$$
$$d_2(\lambda)=(\lambda-c_1)^{l_{21}}(\lambda-c_2)^{l_{22}}\cdots(\lambda-c_r)^{l_{2r}}$$
$$\vdots$$
$$d_s(\lambda)=(\lambda-c_1)^{l_{s1}}(\lambda-c_2)^{l_{s2}}\cdots(\lambda-c_r)^{l_{sr}}.$$

此处，l_{ij} 为非负整数且有

$l_{11}\leqslant l_{21}\leqslant\cdots\leqslant l_{s1}$; $l_{12}\leqslant l_{22}\leqslant\cdots\leqslant l_{s2}$; $\cdots$; $l_{1r}\leqslant l_{2r}\leqslant\cdots\leqslant l_{sr}$。于是幂

$(\lambda-c_1)^{l_{11}},(\lambda-c_2)^{l_{12}}\cdots,(\lambda-c_r)^{l_{1r}}$ 中指数为正的就是$d_1(\lambda)$的初等因子；

$(\lambda-c_1)^{l_{21}},(\lambda-c_2)^{l_{22}},\cdots,(\lambda-c_r)^{l_{2r}}$ 中指数为正的就是 $d_2(\lambda)$的初等因子；

··························

$(\lambda-c_1)^{l_{s1}}$, $(\lambda-c_2)^{l_{s2}},\cdots,(\lambda-c_r)^{l_{sr}}$ 中指数为正的就是 $d_s(\lambda)$的初等因子。

我们把 $d_1(\lambda)$，$d_2(\lambda),\cdots,d_s(\lambda)$的每一个初等因子都叫做 $\boldsymbol{A}(\lambda)$ **的初等因子**，其全体初等因子叫做 $\boldsymbol{A}(\lambda)$ **的初等因子组**。若$A(\lambda)$是A的特征矩阵时，那么 $A(\lambda)=\lambda I-A$ 的初等因子、初等因子组分别叫做**A的初等因子、初等因子组**。

于是对每一个复数矩阵A，把A的特征矩阵 $\lambda I-A$化为定理的形式，求出A的不变因子，通过不变因子的标准分解式，求出A的初等因子组。

2）求若当块。对每一个幂 $(\lambda-c)^m$都与下列形式相当简单的m 阶方阵是互相唯一决定的：

$$N=\begin{pmatrix} c & 1 & 0 & \cdots & 0 & 0\\ 0 & c & 1 & \cdots & 0 & 0\\ 0 & 0 & c & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c & 1\\ 0 & 0 & 0 & \cdots & 0 & c \end{pmatrix}.$$

把矩阵N叫做$(\lambda-c)^m$的若当块。

这样，如果A的初等因子组已经求出，那么对应于每个初等因子，都有唯一确定的若当块：

$$N_{11}=\overset{(\lambda-c_1)^{l_{11}}\downarrow}{\begin{pmatrix} c_1 & 1 & 0 & \cdots & 0 & 0\\ 0 & c_1 & 1 & \cdots & 0 & 0\\ 0 & 0 & c_1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c_1 & 1\\ 0 & 0 & 0 & \cdots & 0 & c_1 \end{pmatrix}}_{(l_{11}\text{阶的})},\quad N_{12}=\overset{(\lambda-c_2)^{l_{12}}\downarrow}{\begin{pmatrix} c_2 & 1 & 0 & \cdots & 0 & 0\\ 0 & c_2 & 1 & \cdots & 0 & 0\\ 0 & 0 & c_2 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c_2 & 1\\ 0 & 0 & 0 & \cdots & 0 & c_2 \end{pmatrix}}_{(l_{12}\text{阶的})},\cdots,$$

$$N_{1r}=\overset{(\lambda-c_r)^{l_{1r}}\downarrow}{\begin{pmatrix} c_r & 1 & 0 & \cdots & 0 & 0\\ 0 & c_r & 1 & \cdots & 0 & 0\\ 0 & 0 & c_r & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & & c_r & 1\\ 0 & 0 & 0 & \cdots & 0 & c_r \end{pmatrix}}_{(l_{1r}\text{阶的})};\quad N_{21}=\overset{(\lambda-c_1)^{l_{21}}\downarrow}{\begin{pmatrix} c_1 & 1 & 0 & \cdots & 0 & 0\\ 0 & c_1 & 1 & \cdots & 0 & 0\\ 0 & 0 & c_1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c_1 & 1\\ 0 & 0 & 0 & \cdots & 0 & c_1 \end{pmatrix}}_{(l_{21}\text{ 阶的})},$$

$$N_{22}=\overset{(\lambda-c_2)^{l_{22}}\downarrow}{\begin{pmatrix} c_2 & 1 & 0 & \cdots & 0 & 0\\ 0 & c_2 & 1 & \cdots & 0 & 0\\ 0 & 0 & c_2 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c_2 & 1\\ 0 & 0 & 0 & \cdots & 0 & c_2 \end{pmatrix}}_{(l_{22}\text{ 阶的})},\cdots,N_{2r}=\overset{(\lambda-c_r)^{l_{2r}}\downarrow}{\begin{pmatrix} c_r & 1 & 0 & \cdots & 0 & 0\\ 0 & c_r & 1 & \cdots & 0 & 0\\ 0 & 0 & c_r & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c_r & 1\\ 0 & 0 & 0 & \cdots & 0 & c_r \end{pmatrix}}_{(l_{2r}\text{阶的})};$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots;$$

$$N_{s1}=\overset{(\lambda-c_1)^{l_{s1}}\downarrow}{\begin{pmatrix} c_1 & 1 & 0 & \cdots & 0 & 0\\ 0 & c_1 & 1 & \cdots & 0 & 0\\ 0 & 0 & c_1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c_1 & 1\\ 0 & 0 & 0 & \cdots & 0 & c_1 \end{pmatrix}}_{(l_{s1}\text{ 阶的})},N_{s2}=\overset{(\lambda-c_2)^{l_{s2}}\downarrow}{\begin{pmatrix} c_2 & 1 & 0 & \cdots & 0 & 0\\ 0 & c_2 & 1 & \cdots & 0 & 0\\ 0 & 0 & c_2 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & c_2 & 1\\ 0 & 0 & 0 & \cdots & 0 & c_2 \end{pmatrix}}_{(l_{s2}\text{ 阶的})},\cdots,$$

$$(\lambda - c_r)^{l_{sr}}$$
$$\downarrow$$
$$N_{sr} = \begin{pmatrix} c_r & 1 & 0 & \cdots & 0 & 0 \\ 0 & c_r & 1 & \cdots & 0 & 0 \\ 0 & 0 & c_r & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & c_r & 1 \\ 0 & 0 & 0 & \cdots & 0 & c_r \end{pmatrix}.$$

(l_{sr} 阶的)

当然，为了整齐，我们写出了一切可能的初等因子及若当块。当某个指数 $l_{ij}=0$时，幂 $(\lambda-c_j)^{l_{ij}}$不是初等因子，自然也就没有相应的若当块。为了说话方便，把上述若当块的全体叫做A的若当块组。

3）组成若当型矩阵。设A的若当块组已经求出如下：

$$N_{11},\cdots,N_{1r};N_{21},\cdots,N_{2r};\cdots;\ N_{s1},\cdots,N_{sr}.$$

于是，我们可以组成一个形式相当简单的方阵如

$$B = \begin{pmatrix} N_{11} & & & & & \\ & \ddots & & & & \\ & & N_{1r} & & & \\ & & & N_{21} & & \\ & & & & \ddots & \\ & & & & & N_{2r} \\ & & & & & & \ddots \\ & & & & & & & N_{s1} \\ & & & & & & & & \ddots \\ & & & & & & & & & N_{sr} \end{pmatrix}.$$

这是一个分块对角形矩阵，它对角线上各小块恰好构成A的若当块组。把B这样的矩阵叫做阵 **A的若当型阵**。

这样，对任一复数域上n阶方阵A都能求出一个若当型矩阵B。我们的结论是：A与它的若当型矩阵B相似：$A\sim B$。

结论的正确性放在下一节去证明。把与A相似的若当型矩阵叫做**A的若当标准形**。

总括起来便得：对复数域上任一n阶方阵A都有若当标准形，而且可以按照上述三个步骤具体地求出A的若当标准形。

下面通过例子来熟悉求若当标准形的方法和步骤.

例 4　求矩阵

$$A=\begin{pmatrix}-1 & -2 & 6\\ -1 & 0 & 3\\ -1 & -1 & 4\end{pmatrix}$$

的若当标准形.

解　这就是例 1 那个矩阵

1）求A的初等因子组. 初等因子组是通过不变因子的标准分解式得来的. 这里A的不变因子已在例 1 中求得为:

$$d_1(\lambda)=1,\ d_2(\lambda)=\lambda-1,\ d_3(\lambda)=\lambda^2-2\lambda+1$$

它们在复数域上的标准分解式为

$$d_1(\lambda)=1,\ d_2(\lambda)=\lambda-1,\ d_3(\lambda)=(\lambda-1)^2.$$

于是, $d_1(\lambda)$ 没有初等因子, $d_2(\lambda)$ 只有一个初等因子 $\lambda-1$, $d_3(\lambda)$ 也只有一个初等因子 $(\lambda-1)^2$. 所以, A的初等因子组为

$$\lambda-1,\ (\lambda-1)^2.$$

2）求若当块.

初等因子 $\lambda-1$ 的若当块为 $N_1=(1)$;

初等因子 $(\lambda-1)^2$ 的若当块为 $N_2=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$.

即A的若当块组为

$$(1),\ \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}.$$

3）组成若当型阵. 把A的若当块组摆在对角线上, 就得到A的若当型阵

$$B=\begin{pmatrix}1 & & \\ & 1 & 1\\ & 0 & 1\end{pmatrix}.$$

于是, B 就是A的若当标准形, 即

$$A\sim B=\begin{pmatrix}1 & & \\ & 1 & 1\\ & 0 & 1\end{pmatrix}.$$

例 5 求矩阵

$$A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$$

的若当标准形。

解 1）求初等因子组．在例 2 中已求出 A 的不变因子为

$$d_1(\lambda)=1,\ d_2(\lambda)=\lambda+1,\ d_3(\lambda)=\lambda^2-4\lambda-5.$$

于是 A 的初等因子组为 $\lambda+1,\ \lambda+1,\lambda-5$.

2）求若当块。

$\lambda+1$ 的若当块为 $N_1=(-1)$，$\lambda+1$ 的若当块为 $N_2=(-1)$，$\lambda-5$ 的若当块为 $N_3=(5)$.

3）组成若当型矩阵．以若当块 N_1,N_2,N_3 为对角线上的小块组成的若当型矩阵：

$$B=\begin{pmatrix}N_1&&\\&N_2&\\&&N_3\end{pmatrix}=\begin{pmatrix}-1&&\\&-1&\\&&5\end{pmatrix}.$$

这个若当型阵 B 就是 A 的若当标准形，即

$$A\sim B=\begin{pmatrix}-1&&\\&-1&\\&&5\end{pmatrix}.$$

例 6 求矩阵

$$A=\begin{pmatrix}1&2&0\\-1&1&-1\\0&-2&1\end{pmatrix}$$

的若当标准形。

解 1）求初等因子组．已在例 3 中求出 A 的不变因子为

$$d_1(\lambda)=1,d_2(\lambda)=1,\ d_3(\lambda)=(\lambda-1)^3.$$

所以 A 的初等因子组为 $(\lambda-1)^3$.

2）求若当块．

$$(\lambda-1)^3 \text{ 的若当块为 } N_1=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}.$$

即A的若当块组，只此一个块。

3）组成若当型矩阵。因为若当块组只有一个阵N_1，所以由若当块组组成的若当型矩阵就是N_1。因此，A的若当标准形就是N_1，即

$$A\sim N_1=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}.$$

例7　求矩阵

$$A=\begin{pmatrix}0&0&0&2\\0&0&1&0\\0&-1&0&0\\1&0&0&0\end{pmatrix}$$

的有理标准形及若当标准形。

解1）求A的不变因子和初等因子组。

$$\lambda I-A=\begin{pmatrix}\lambda&0&0&-2\\0&\lambda&-1&0\\0&1&\lambda&0\\-1&0&0&\lambda\end{pmatrix}\xrightarrow{P_{14}(\lambda)}\begin{pmatrix}0&0&0&\lambda^2-2\\0&\lambda&-1&0\\0&1&\lambda&0\\-1&0&0&\lambda\end{pmatrix}$$

$$\xrightarrow[P_{14}(\lambda)]{}\begin{pmatrix}0&0&0&\lambda^2-2\\0&\lambda&-1&0\\0&1&\lambda&0\\-1&0&0&0\end{pmatrix}\xrightarrow{D_4(-1),C_{14}}\begin{pmatrix}1&0&0&0\\0&\lambda&-1&0\\0&1&\lambda&0\\0&0&0&\lambda^2-2\end{pmatrix}$$

$$\xrightarrow{P_{23}(-\lambda)}\begin{pmatrix}1&0&0&0\\0&0&-\lambda^2-1&0\\0&1&\lambda&0\\0&0&0&\lambda^2-2\end{pmatrix}\xrightarrow{P_{23}(-\lambda)}\begin{pmatrix}1&0&0&0\\0&0&-\lambda^2-1&0\\0&1&0&0\\0&0&0&\lambda^2-2\end{pmatrix}$$

$$\xrightarrow{C_{23}}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-\lambda^2-1&0\\0&0&0&\lambda^2-2\end{pmatrix}\xrightarrow{P_{34}(1)}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-\lambda^2-1&\lambda^2-2\\0&0&0&\lambda^2-2\end{pmatrix}$$

$$\xrightarrow[P_{43}(1)]{}\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-3&\lambda^2-2\\0&0&\lambda^2-2&\lambda^2-2\end{pmatrix}\xrightarrow{P_{43}\left(\frac{1}{3}(\lambda^2-2)\right)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & \lambda^2-2 \\ 0 & 0 & 0 & \frac{1}{3}(\lambda^2-2)^2+(\lambda^2-2) \end{pmatrix} \xrightarrow{P_{34}\left(\frac{1}{3}(\lambda^2-2)\right)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & \frac{1}{3}(\lambda^2-2)^2+(\lambda^2-2) \end{pmatrix} \xrightarrow{D_3\left(-\frac{1}{3}\right)}$$

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \frac{1}{3}(\lambda^2-2)(\lambda^2+1) \end{pmatrix} \xrightarrow{D_4(3)} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & (\lambda^2-2)(\lambda^2+1) \end{pmatrix}.$$

到这一步已求出A的不变因子为

$$1, 1, 1, (\lambda^2-2)(\lambda^2+1)=\lambda^4-\lambda^2-2,$$

初等因子组为

$$\lambda-\sqrt{2}, \lambda+\sqrt{2}, \lambda-i, \lambda+i.$$

2）求伴侣矩阵和若当块.

因为非零次的不变因子只有一个 $\lambda^4-\lambda^2-2$，所以伴侣矩阵也只有一个：

$$\begin{pmatrix} 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

因为初等因子共有四个，所以若当块也有四个：

$(\sqrt{2})$, $(-\sqrt{2})$，(i)，$(-i)$.

3）组成有理型阵和若当型阵.

A的有理型阵为

$$B=\begin{pmatrix} 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

A的若当型阵为

$$C=\begin{pmatrix}\sqrt{2} & & & \\ & -\sqrt{2} & & \\ & & i & \\ & & & -i\end{pmatrix}.$$

故A的有理标准形为B，若当标准形为C。

练　习　七

1．求下列矩阵的有理标准形：

(1) $\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$，(2) $\begin{pmatrix}1 & 2 & 0\\ 0 & 2 & 0\\ -2 & -2 & -1\end{pmatrix}$，

(3) $\begin{pmatrix}3 & 1 & 0 & 0\\ -4 & -1 & 0 & 0\\ 7 & 1 & 2 & 1\\ -17 & -6 & -1 & 0\end{pmatrix}$。

2．求下列矩阵的若当标准形：

(1) $\begin{pmatrix}1 & 2 & 0\\ 0 & 2 & 0\\ -2 & -2 & -1\end{pmatrix}$，(2) $\begin{pmatrix}4 & 6 & 0\\ -3 & -5 & 0\\ -3 & -6 & 1\end{pmatrix}$，

(3) $\begin{pmatrix}3 & 0 & 8\\ 3 & -1 & 6\\ -2 & 0 & -5\end{pmatrix}$，(4) $\begin{pmatrix}3 & 1 & 0 & 0\\ -4 & -1 & 0 & 0\\ 7 & 1 & 2 & 1\\ -17 & -6 & -1 & 0\end{pmatrix}$。

3．证明

1) $A=\begin{pmatrix}0 & 0 & 1\\ 1 & 0 & -3\\ 0 & 1 & 3\end{pmatrix}$

的非零次的不变因子恰为 $\lambda^3-3\lambda^2+3\lambda-1$；

2)

$$B=\begin{pmatrix}0 & 0 & 0 & \cdots & 0 & -a_m\\ 1 & 0 & 0 & \cdots & 0 & -a_{m-1}\\ 0 & 1 & 0 & \cdots & 0 & -a_{m-2}\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & 0 & -a_2\\ 0 & 0 & 0 & \cdots & 1 & -a_1\end{pmatrix}$$

的非零次的不变因子恰为

$$\lambda^m + a_1\lambda^{m-1} + a_2\lambda^{m-2} + \cdots + a_{m-1}\lambda + a_m.$$

4. 证明

1） $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$

的初等因子组恰为 $(\lambda-2)^3$；

2） m阶矩阵

$$B = \begin{pmatrix} \rho & 1 & 0 & \cdots & 0 & 0 \\ 0 & \rho & 1 & \cdots & 0 & 0 \\ 0 & 0 & \rho & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \rho & 1 \\ 0 & 0 & 0 & \cdots & 0 & \rho \end{pmatrix}$$

的初等因子组恰为 $(\lambda-\rho)^m$。

5. 设

$$A = \begin{pmatrix} 1 & 4 & 2 \\ 0 & -3 & 4 \\ 0 & 4 & 3 \end{pmatrix}$$

求 A^k。

§8 λ—矩阵在初等变换之下的化简　相似定理

上一节介绍了相似矩阵的两种标准形及其求法。当时有两个关键问题留在这一节来解决：

一个问题是，任意 λ—矩阵 $A(\lambda)$ $(\neq 0)$ 经若干次初等变换都能化简成形式：

$$\begin{pmatrix} d_1(\lambda) & & & & \\ & \ddots & & & \\ & & d_s(\lambda) & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

其中 $d_i(\lambda)$ 都是首项系数为 1 的多项式，并且依次能整除；

$$d_1(\lambda) \mid d_2(\lambda) \mid \cdots \mid d_s(\lambda).$$

另一个问题是，对任一方阵A都有它自己的不变因子和初等因子。由不变因子确定的唯一的有理型矩阵

$$N=\begin{pmatrix} N_1 & & & \\ & N_2 & & \\ & & \ddots & \\ & & & N_r \end{pmatrix},$$

其中 $N_1, N_2, \cdots, N_r$ 恰是A的非零次不变因子的伴侣矩阵；由A的初等因子组确定的一个若当型矩阵

$$J=\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{pmatrix},$$

其中 $J_1, J_2, \cdots, J_t$ 恰是A的一切初等因子的若当块。那么

$$A \sim N, \quad A \sim J.$$

本节主要就是证明这两个问题。首先做一些叙述上的准备。

在上节已经说过一般λ—矩阵的概念以及与其有关的问题：

1）λ—矩阵的初等变换；

2）λ—矩阵的行列式、子式与秩数；

3）λ—矩阵的加法、减法与乘法。

这里再明确补充以下几点：

4） 两个λ-矩阵$A(\lambda)$，$B(\lambda)$，如果用初等变换能把$A(\lambda)$变为$B(\lambda)$，则称$A(\lambda)$与$B(\lambda)$相抵，记作$A(\lambda) \to B(\lambda)$。像相合、相似关系一样，容易验证：两个$\lambda$—矩阵的相抵关系是等价关系，这个验证留给读者。

5） 初等λ—矩阵：

倍法阵

$$D_i(d)=\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}(i) \quad (d \neq 0),$$

消法阵

$$P_{ij}(k(\lambda))=\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1\cdots k(\lambda)&&\\&&\ddots&\vdots&\\&&&1&\\&&&&\ddots\\&&&&&1\end{pmatrix}\begin{matrix}\\\\(i)\\\\\\\\\\\end{matrix}$$
$$(j)$$

或

$$P_{ji}(k(\lambda))=\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&\vdots&\ddots&\\&&k(\lambda)\cdots1&&\\&&&&\ddots\\&&(i)&&&1\end{pmatrix}(j)。$$

换法阵

$$C_{ij}=\begin{pmatrix}1&&&&&&&\\&\ddots&&&&&&\\&&1&&&&&\\&&&0&\cdots\cdots&1&&\\&&&\vdots&1&\vdots&&\\&&&\vdots&\ddots&\vdots&&\\&&&\vdots&1&\vdots&&\\&&&1&\cdots\cdots&0&&\\&&&&&&1&\\&&&&&&&\ddots\\&&&&&&&&1\end{pmatrix}\begin{matrix}\\\\\\(i)\\\\\\\\(j)\\\\\\\\\end{matrix}。$$

容易验证，用初等矩阵去左（右）乘 $A(\lambda)$，恰好相当于用相应的初等变换去变 $A(\lambda)$的行（列）。从而有：

$A(\lambda)\to B(\lambda)$ 必要而且只要

$$B(\lambda)=P_s(\lambda)\cdots P_1(\lambda)A(\lambda)Q_1(\lambda)\cdots Q_t(\lambda),$$

其中 $P_i(\lambda),Q_i(\lambda)$皆为初等矩阵.

6） λ—矩阵 $A(\lambda)$ 的行列式，子式一般来说是一个多项式. $A(\lambda)$经初等变换之后，行列式、子式可能有所变化，但 $A(\lambda)$的秩数是不变的，这是下面的命题所要回答的。

命题 1　λ—矩阵 $A(\lambda)$ 经若干次初等变换后，秩数不变.

证明　显然倍法变换不改变 λ—矩阵的秩数，故只须证消法变换不改变 λ—矩阵的秩数. 我们只证对行的消法变换命题 1 成立，因为列的情形可同理证之. 设

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1}(\lambda) & a_{i2}(\lambda) & \cdots & a_{in}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ a_{j1}(\lambda) & a_{j2}(\lambda) & \cdots & a_{jn}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots & a_{mn}(\lambda) \end{pmatrix}$$

经消法变换后，变为

$$B(\lambda)=$$

$$\begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots a_{1n}(\lambda) \\ \cdots & \cdots & \cdots \\ a_{i1}(\lambda) & a_{i2}(\lambda) & \cdots a_{in}(\lambda) \\ \cdots & \cdots & \cdots \\ a_{j1}(\lambda)+k(\lambda)a_{i1}(\lambda) & a_{j2}(\lambda)+k(\lambda)a_{i2}(\lambda) & \cdots a_{jn}(\lambda)+k(\lambda)a_{in}(\lambda) \\ \cdots & \cdots & \cdots \\ a_{m1}(\lambda) & a_{m2}(\lambda) & \cdots a_{mn}(\lambda) \end{pmatrix},$$

且设 $A(\lambda)$的秩数为 r，$B(\lambda)$ 的秩数为 s，要证的是 $r=s$.

在 $B(\lambda)$ 中任取一个 $r+1$ 阶子式，记为 $D_{r+1}(\lambda)$. 则此子式可分为三种情况：

1）当$D_{r+1}(\lambda)$不含 $B(\lambda)$的第 j 行时，那么 $D_{r+1}(\lambda)$ 与 $A(\lambda)$ 中的某一子式完全一样；

2）当 $D_{r+1}(\lambda)$ 含 $B(\lambda)$的第 j 行，也含第 i 行时，那么 $D_{r+1}(\lambda)$ 可分为两个 $r+1$ 阶子式：

$$D_{r+1}(\lambda)=$$

$$\begin{vmatrix} \cdots & \cdots & \cdots \\ a_{il_1}(\lambda) & a_{il_2}(\lambda) & \cdots\ a_{il_{r+1}}(\lambda) \\ \cdots & \cdots & \cdots \\ a_{jl_1}(\lambda)+k(\lambda)a_{il_1}(\lambda) & a_{jl_2}(\lambda)+k(\lambda)a_{il_2}(\lambda) & \cdots a_{jl_{r+1}}(\lambda)+k(\lambda)a_{il_{r+1}}(\lambda) \\ \cdots & \cdots & \cdots \end{vmatrix}$$

$$=\begin{vmatrix} \cdots\cdots\cdots\cdots \\ a_{il_1}(\lambda) & a_{il_2}(\lambda) & \cdots a_{il_{r+1}}(\lambda) \\ \cdots\cdots\cdots\cdots \\ a_{jl_1}(\lambda) & a_{jl_2}(\lambda) & \cdots a_{jl_{r+1}}(\lambda) \\ \cdots\cdots\cdots\cdots \end{vmatrix} + \begin{vmatrix} \cdots\cdots\cdots\cdots \\ a_{il_1}(\lambda) & a_{il_2}(\lambda) & \cdots a_{il_{r+1}}(\lambda) \\ \cdots\cdots\cdots\cdots \\ k(\lambda)a_{il_1}(\lambda) & k(\lambda)a_{il_2}(\lambda) & \cdots k(\lambda)a_{il_{r+1}}(\lambda) \\ \cdots\cdots\cdots\cdots \end{vmatrix}$$

其中之一是 $A(\lambda)$的一个子式，另一个为0，因此该子式就是 $A(\lambda)$中的一个子式；

3）当 $D_{r+1}(\lambda)$ 含 $B(\lambda)$ 的第 j 行不含第 i 行时，$D_{r+1}(\lambda)$可分为两个 $r+1$ 阶子式：

$D_{r+1}(\lambda)=$

$$\begin{vmatrix} \cdots\cdots\cdots\cdots \\ a_{jl_1}(\lambda)+k(\lambda)a_{il_1}(\lambda) & a_{jl_2}(\lambda)+k(\lambda)a_{il_2}(\lambda) & \cdots a_{jl_{r+1}}(\lambda)+k(\lambda)a_{il_{r+1}}(\lambda) \\ \cdots\cdots\cdots\cdots \end{vmatrix}$$

$$=\begin{vmatrix} \cdots\cdots\cdots\cdots \\ a_{jl_1}(\lambda) & a_{jl_2}(\lambda) & \cdots a_{jl_{r+1}}(\lambda) \\ \cdots\cdots\cdots\cdots \end{vmatrix} + k(\lambda)\begin{vmatrix} \cdots\cdots\cdots\cdots \\ a_{il_1}(\lambda) & a_{il_2}(\lambda) & \cdots a_{il_{r+1}}(\lambda) \\ \cdots\cdots\cdots\cdots \end{vmatrix}$$

其中之一是 $A(\lambda)$ 的一个子式，另一个是 $A(\lambda)$的一个子式的$k(\lambda)$倍或$-k(\lambda)$倍。

由于 $A(\lambda)$ 的秩数是 r，$A(\lambda)$ 中任一$r+1$阶子式都为0，于是不管上述那种情况，$B(\lambda)$ 中的任一 $r+1$ 阶子式必都为 0。因此，$B(\lambda)$ 的秩数 $s\leqslant r$。

上边已经证明了，一个 λ—矩阵经行的消法变换后，秩数不增加。

由于 $B(\lambda)$ 又可经消法变换变为 $A(\lambda)$（这只要把 $B(\lambda)$ 的第 i 行 $-k(\lambda)$倍加到第 j 行上），于是 $A(\lambda)$ 的秩数 $r\leqslant s$，故 $r=s$，命题证完。

7）n 阶方阵 $A(\lambda)$ 叫做可逆的，如果存在 $B(\lambda)$，使

$$A(\lambda)B(\lambda)=I_n=B(\lambda)A(\lambda),$$

其中 I_n 为通常的单位阵。 于是，有

命题 2　$A(\lambda)$ 是可逆的充分必要条件是 $A(\lambda)$ 的行列式是一个非零常数。

证明　必要性．设 $A(\lambda)$是可逆的，则存在 $B(\lambda)$，使

$$A(\lambda)B(\lambda)=I_n.$$

于是，一方面

$$\det(A(\lambda)B(\lambda))=\det A(\lambda)\cdot\det B(\lambda).$$

另一方面

$$\det(A(\lambda)B(\lambda))=\det I_n=1.$$

故

$$\det A(\lambda)\cdot\det B(\lambda)=1.$$

因此，必有

$$\det A(\lambda)=d\neq 0,\text{ 即 } d \text{ 为非零常数.}$$

充分性．若$\det A(\lambda)=d\neq 0$（d 为非零常数），作 $A(\lambda)$ 的伴随矩阵 $A^*(\lambda)$．令 $B(\lambda)=\frac{1}{d}A^*(\lambda)$，则有

$$A(\lambda)B(\lambda)=\frac{1}{d}A(\lambda)A^*(\lambda)=I_n=\frac{1}{d}A^*(\lambda)A(\lambda)=B(\lambda)A(\lambda).$$

因此 $A(\lambda)$是可逆的，证完．

由命题 2 显然有，初等 λ—阵都是可逆的；可逆 λ—矩阵之积仍是可逆的．

有了上述准备，下面就来解决当初提出的两个遗留问题．

定理 1　设$A(\lambda)$为任一非零的 λ—矩阵．用初等变换必能把 $A(\lambda)$ 化简成对角形阵：

$$\begin{pmatrix} d_1(\lambda) & & & & & \\ & d_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & d_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix} \quad (1)$$

其中 $d_i(\lambda)$ 皆为首系数是 1 的多项式，并且依次能整除：$d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_r(\lambda)$．

证明　考虑一切与 $A(\lambda)$ 相抵的 λ—矩阵．在这些矩阵当中，

左上角元素，即(1,1)位置的元素必有次数最低的．于是不妨认为$A(\lambda)$就是这样的．即$a_{11}(\lambda)$是与$A(\lambda)$相抵的一切λ—矩阵当中左上角元素次数为最低的一个（同时可以认为其首系数为1）．于是我们可以断定：$a_{11}(\lambda)$能整除$A(\lambda)$第一行(列)的每一个元素．如若不然，必有

$$a_{1j}=a_{11}(\lambda)q(\lambda)+r(\lambda),\quad r(\lambda)\text{的次数}<a_{11}(\lambda)\text{的次数．这样}$$

$$A(\lambda)=\begin{pmatrix} a_{11}(\lambda)\cdots & [a_{11}(\lambda)q(\lambda)+r(\lambda)] & \cdots a_{1n}(\lambda) \\ a_{21}(\lambda)\cdots & a_{2j}(\lambda) & \cdots a_{2n}(\lambda) \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{m1}(\lambda)\cdots & a_{mj}(\lambda) & \cdots a_{mn}(\lambda) \end{pmatrix}.$$

经用$-q(\lambda)$乘$A(\lambda)$的第一列，加到第j列上，然后交换$1,j$两列可使$A(\lambda)$相抵于形式如下的矩阵

$$\begin{pmatrix} r(\lambda) & \cdots \\ \cdots & \ddots \end{pmatrix}.$$

这和$a_{11}(\lambda)$是与$A(\lambda)$相抵矩阵中左上角元素次数最低的前提相矛盾．因此必有

$$a_{11}(\lambda)\,|\,a_{1j}(\lambda),\quad j=2,\cdots,n,$$

同理也有

$$a_{11}(\lambda)\,|\,a_{i1}(\lambda),\quad i=2,\cdots,m.$$

这样，用适当的初等变换，就可使A相抵于

$$B(\lambda)=\begin{pmatrix} a_{11}(\lambda) & 0 \cdots 0 \\ 0 & \\ \vdots & A_1(\lambda) \\ 0 & \end{pmatrix}.$$

如果$A_1(\lambda)=0$，$B(\lambda)$即合所求．如果$A_1(\lambda)\neq 0$，重复以上过程可知

$$A_1(\lambda)\rightarrow\begin{pmatrix} b_{22}(\lambda) & 0 \cdots 0 \\ 0 & \\ \vdots & A_2(\lambda) \\ 0 & \end{pmatrix}.$$

其中，$b_{22}(\lambda)$也具有次数最低的性质。

从而

$$A(\lambda) \to \begin{pmatrix} a_{11}(\lambda) & & \\ & b_{22}(\lambda) & \\ & & A_2(\lambda) \end{pmatrix}.$$

如此继续下去，必有

$$A(\lambda) \to \begin{pmatrix} d_1(\lambda) & & & & & & \\ & d_2(\lambda) & & & & & \\ & & \ddots & & & & \\ & & & d_r(\lambda) & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

其中 $d_1(\lambda)$，$d_2(\lambda),\cdots,d_r(\lambda)$都是首系数为 1 的多项式。

根据 $d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$ 都具有"次数最低"的性质可知它们是依次能整除的：$d_1(\lambda)|d_2(\lambda)|\cdots|d_r(\lambda)$。这是因为如若不然，必有 $d_i(\lambda)\nmid d_{i+1}(\lambda)$，那么

$$d_{i+1}(\lambda)=d_i(\lambda)q_1(\lambda)+r_1(\lambda)\text{，其中 } \deg r_1(x)<\deg d_i(\lambda).$$

这时

$$A(\lambda) \to \begin{pmatrix} d_1(\lambda) & & & & & & & \\ & \ddots & & & & & & \\ & & d_i(\lambda) & & & & & \\ & & & [d_i(\lambda)q_1(\lambda)+r_1(\lambda)] & & & & \\ & & & & \ddots & & & \\ & & & & & d_r(\lambda) & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}$$

$$\to \begin{pmatrix} d_1(\lambda) & & & & & & \\ & \ddots & & & & & \\ & & d_i(\lambda) & [d_i(\lambda)q_1(\lambda)+r_1(\lambda)] & & & \\ & & & [d_i(\lambda)q_1(\lambda)+r_1(\lambda)] & & & \\ & & & & \ddots & & \\ & & & & & d_r(\lambda) & \\ & & & & & & 0 \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}$$

$$\rightarrow\begin{pmatrix} d_1(\lambda) & & & & & & & \\ & \ddots & & & & & & \\ & & d_i(\lambda) & r_1(\lambda) & & & & \\ & & & d_{i+1}(\lambda) & & & & \\ & & & & \ddots & & & \\ & & & & & d_r(\lambda) & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}$$

$$\rightarrow\begin{pmatrix} d_1(\lambda) & & & & & & & \\ & \ddots & & & & & & \\ & & r_1(\lambda) & d_i(\lambda) & & & & \\ & & d_{i+1}(\lambda) & 0 & & & & \\ & & & & \ddots & & & \\ & & & & & d_r(\lambda) & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}$$

这样出现了与$A_{i-1}(\lambda)$相抵且(1,1)位置元素$r_1(\lambda)$的次数$<d_i(\lambda)$的次数，这是一个矛盾．故定理得证．

推论1 $A(\lambda)$ **可逆的充分必要条件是**

$$A(\lambda)=T_1T_2\cdots T_m,$$

其中 T_i **为初等** λ**—矩阵．**

证明 充分性是明显的，我们主要证必要性．设 $A(\lambda)$ 可逆，据定理1 $A(\lambda)$相抵于

$$B(\lambda)=\begin{pmatrix} d_1(\lambda) & & & \\ & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{pmatrix},$$

其中$d_i(\lambda)\neq 0$，$i=1,2,\cdots,n$．也就是

$$B(\lambda)=P_t(\lambda)\cdots P_1(\lambda)A(\lambda)Q_1(\lambda)\cdots Q_t(\lambda)$$

其中 $P_i(\lambda),Q_j(\lambda)$ 都是初等 λ—矩阵．由 $P_i(\lambda),Q_j(\lambda),A(\lambda)$ 都是可逆 λ—矩阵，则它们之积也是可逆 λ—矩阵，于是 $B(\lambda)$ 是可逆 λ—矩阵，据命题1有

$$\det B(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)=1 \text{（因 } d_i(\lambda)\text{ 首系数为 1）},$$

因此，每个 $d_i(\lambda)=1$，$i=1,2,\cdots,n$。故 $B(\lambda)=I_n$，从而

$$P_s(\lambda)\cdots P_1(\lambda)A(\lambda)Q_1(\lambda)\cdots Q_t(\lambda)=I_n$$

即

$$A(\lambda)=P_1^{-1}(\lambda)\cdots P_s^{-1}(\lambda)Q_t^{-1}(\lambda)\cdots Q_1^{-1}(\lambda)$$
$$=T_1\cdots T_m。$$ 证完。

由推论 1 立即得到

推论 2 $\boldsymbol{A}(\lambda)\to\boldsymbol{B}(\lambda)$ **充分必要条件是：存在可逆 λ—阵 $\boldsymbol{P}(\lambda)$，$\boldsymbol{Q}(\lambda)$，使**

$$\boldsymbol{B}(\lambda)=\boldsymbol{P}(\lambda)\boldsymbol{A}(\lambda)\boldsymbol{Q}(\lambda)。$$

由定理 1，从 $A(\lambda)$ 化简得到的对角形 λ—阵（1）叫做 $A(\lambda)$ 在相抵之下的标准形。我们将指出，λ—矩阵在相抵之下的标准形是唯一的。为此，引入行列式因子的概念。

设 $A(\lambda)$ 的秩为 r。于是对于任意一个 s：$1\leqslant s\leqslant r$，$A(\lambda)$的一切 s 阶子式是一组不全为 0 的多项式，用$D_s(\lambda)$ 表示这组多项式的最高公因式，从而得到由 $A(\lambda)$所唯一确定的一组非零的多项式：

$$D_1(\lambda),D_2(\lambda),\cdots,D_r(\lambda)。$$

把上边的每一个多项式 $D_s(\lambda)$（$1\leqslant s\leqslant r$）叫做 **$A(\lambda)$的（s 阶的）行列式因子**。

据 $D_s(\lambda)$ 的定义和行列式的展开定理，显然这组行列式因子有依次能整除的性质：

$$D_1(\lambda)\,|D_2(\lambda)|\cdots|D_r(\lambda)。$$

行列式因子的重要性在于它们在初等变换之下是不变的，即有

定理 2 **若 $\boldsymbol{A}(\lambda)$与 $\boldsymbol{B}(\lambda)$相抵，则 $A(\lambda)$与 $\boldsymbol{B}(\lambda)$有相同的行列式因子。**

证明 $A(\lambda)$ 与 $B(\lambda)$ 相抵，也就是$A(\lambda)$经若干次初等变换得到 $B(\lambda)$。故该定理就是要证 $A(\lambda)$ 经若干次初等变换后，行列式因子不变。于是只须证明经一次初等变换行列式因子不变即可。

首先，看用一次倍法变换从 $A(\lambda)$得到$B(\lambda)$的情形。这时$B(\lambda)$任一 s 阶子式就是$A(\lambda)$ 的一个s 阶子式或者是 $A(\lambda)$的一个 s 阶

子式的非零常数倍。于是，$B(\lambda)$与 $A(\lambda)$的 s 阶行列式因子是相同的，$1\leqslant s\leqslant r$。

其次，考察用一次消法变换从 $A(\lambda)$得到$B(\lambda)$ 的情形。这时从命题 1 的证明可以看出，$B(\lambda)$ 的任一 s 阶子式都能表成

$$M(\lambda)+k(\lambda)N(\lambda)$$

的形式，其中 $M(\lambda),N(\lambda)$都是 $A(\lambda)$的 s 阶子式或 $N(\lambda)$与$A(\lambda)$的某 s 阶子式相差 -1 倍。于是，$A(\lambda)$ 的 s 阶行列式因子必整除 $B(\lambda)$的所有 s 阶子式，从而整除 $B(\lambda)$ 的 s 阶行列式因子；反之，因 $B(\lambda)$经过一次消法变换可得到$A(\lambda)$，所以 $B(\lambda)$的 s 阶行列式因子必整除 $A(\lambda)$的 s 阶行列式因子。故 $B(\lambda)$与 $A(\lambda)$ 的 s 阶行列式因子也是相同的，$1\leqslant s\leqslant r$。

总之，$A(\lambda)$ 经初等变换后，它的各阶行列式因子都不变，因此定理得证。

利用行列式因子在初等变换之下的不变性很容易指出标准形的唯一性，这只须把 $A(\lambda)$ 的标准形与行列式因子联系起来就行了。

定理 3　λ—矩阵 $A(\lambda)$ 在相抵之下的标准形是唯一的。

证明　设 $A(\lambda)$ 的标准形如（1）。于是对角形阵（1）与 $A(\lambda)$ 有相同的行列式因子：$D_1(\lambda)$，$D_2(\lambda),\cdots,D_r(\lambda)$。

我们可以由对角形阵（1）具体计算出各阶行列式因子。事实上，对角形阵（1）的一切非零的一阶子式就是

$$d_1(\lambda),\ d_2(\lambda),\cdots,d_r(\lambda).$$

所以一阶行列式因子 $D_1(\lambda)=d_1(\lambda)$；

对角形阵（1）的一切非零的二阶子式就是

$$d_1(\lambda)d_2(\lambda),d_1(\lambda)d_3(\lambda),\cdots,d_1(\lambda)d_r(\lambda),d_2(\lambda)d_3(\lambda),\\ \cdots,d_{r-1}(\lambda)d_r(\lambda).$$

所以二阶行列式因子$D_2(\lambda)=d_1(\lambda)d_2(\lambda)$；这样下去，对角形阵（1）的一切非零的 r 阶子式就是

$$d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda).$$

所以 r 阶行列式因子$D_r(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda)$。

由此可得

$$d_1(\lambda)=D_1(\lambda),d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)},\ \cdots,\ d_r(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}. \qquad (2)$$

上述（2）式表明：$A(\lambda)$ 的标准形（1）对角线上的非零元素 $d_1(\lambda)$，$d_2(\lambda),\cdots,d_r(\lambda)$由 $A(\lambda)$ 的各阶行列式因子所唯一确定，而行列式因子在初等变换之下是不变的，从而$d_1(\lambda)$，$d_2(\lambda),\cdots$，$d_r(\lambda)$在初等变换之下也是不变的．这就证明了 $A(\lambda)$ 在相抵之下的标准形恰有一个．

把标准形（1）对角线上的非零元素 $d_1(\lambda),d_2(\lambda),\cdots,d_r(\lambda)$ 叫做 $A(\lambda)$的不变因子是恰当的．

既然定理1和定理2保证对任一 λ—矩阵 $A(\lambda)$ 都能具体求出不变因子，而且是唯一的，这也就求出了初等因子，而且初等因子也是由 $A(\lambda)$唯一决定．

这样，不变因子、初等因子组．行列式因子三者都是 λ—矩阵在初等变换之下的不变量．明确地说就是：

两个同秩 λ—矩阵$\boldsymbol{A}(\lambda)$ 与$\boldsymbol{B}(\lambda)$ 相抵的充分必要条件是 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 的不变因子相同；

两个同秩 λ—矩阵 $\boldsymbol{A}(\lambda)$与 $\boldsymbol{B}(\lambda)$相抵的充分必要条件是 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$ 的初等因子组相同；

两个同秩 λ—矩阵 $\boldsymbol{A}(\lambda)$ 与 $\boldsymbol{B}(\lambda)$相抵的充分必要条件是$\boldsymbol{A}(\lambda)$与$\boldsymbol{B}(\lambda)$ 的行列式因子相同．

在上节已给出由A的不变因子来求A的有理标准形；由A的不变因子，再求初等因子组来求A的若当标准形．就是说，求A的有理标准形、若当标准形都需要求出A的不变因子．而如何求A的不变因子呢？到现在为止，有两个办法：

1）根据定义，就是将A的特征矩阵 $\lambda I-A$用初等变换化为标准形，标准形的对角线上的多项式就是A的不变因子；

2）由定理3的证明过程可以看出，先求A的特征矩阵 $\lambda I-A$ 的行列式因子$D_1(\lambda),D_2(\lambda),\cdots,D_n(\lambda)$，然后令

$$d_1(\lambda)=D_1(\lambda),d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)},\cdots,\quad d_n(\lambda)=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}.$$

则 $d_1(\lambda),d_2(\lambda),\cdots,d_n(\lambda)$ 就是A的不变因子。

以上两个办法就是通过化标准形求A的不变因子；通过求行列式因子来求不变因子。我们知道一个 λ—矩阵化成标准形的要求是比较高的，如果 λ—矩阵的行列数稍高一些，阵中的元素稍复杂一些，那么化起来就比较麻烦、比较困难。而一个 λ—矩阵比较特殊一些时求行列式因子还不难，但稍一般一些就不容易了，甚至也比较难。下面再介绍一种求 λ—矩阵不变因子和初等因子的办法，这个办法有时显得方便一些。为此

首先讨论不变因子与初等因子之关系。根据初等因子的定义，我们已经知道，由不变因子可决定初等因子。反之，若知道 λ—矩阵的秩和初等因子组，那么也能决定不变因子。这只要应用不变因子所适合的整除的性质：$d_1(\lambda)\,|d_2(\lambda)\,|\cdots|\,d_r(\lambda)$ 和初等因子组的定义就可以做到。这因为，一是由于初等因子组中每个成员都是某个不变因子 $d_i(\lambda)$ 的因子；二是由于不变因子依次整除的性质知，最后一个不变因子$d_r(\lambda)$ 包含了初等因子组中所有根互异的因子，根相同的因子取方指数最高的那一个。而 $d_{r-1}(\lambda)$ 是初等因子组中，除了 $d_r(\lambda)$ 所含的因子外，包含剩下的所有根互异的因子，根相同的因子取方指数最高的那一个。这样继续下去，最后就决定了不变因子 $d_1(\lambda),\cdots d_r(\lambda)$。当然可能到某步，比如到 $d_k(\lambda)$，初等因子组中的因子全取完了，这时 $d_k(\lambda)=1,d_{k-1}(\lambda)=1,\cdots,d_1(\lambda)=1$。

例如，设 λ—矩阵的秩为 5，初等因子组为

$$(\lambda-1)^2,(\lambda-1)^2,\ (\lambda-1)^2,\ (\lambda+1),\ (\lambda+1)^2,(\lambda-i),(\lambda+i).$$

于是 $A(\lambda)$的不变因子为

$$d_5(\lambda)=(\lambda-1)^2(\lambda+1)^2(\lambda-i)(\lambda+i)=\lambda^6-\lambda^4-\lambda^2+1,$$

$$d_4(\lambda)=(\lambda-1)^2(\lambda+1)=\lambda^3-\lambda^2-\lambda+1,$$

$$d_3(\lambda)=(\lambda-1)^2=\lambda^2-2\lambda+1,$$

$$d_2(\lambda)=1,$$

$d_1(\lambda)=1$。

其次讨论求初等因子组的方法，我们给出

定理 4 如果 $B(\lambda)$是一个 m行 n列的秩为 r的对角形 λ—矩阵：

$$B(\lambda)=\begin{pmatrix} h_1(\lambda) & & & & & \\ & h_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & h_r(\lambda) & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix},$$

而 $h_i(\lambda)$的标准分解式（不妨认为 $h_i(\lambda)$ 首系数为1）为

$$h_i(\lambda)=(\lambda-\lambda_1)^{m_i}(\lambda-\lambda_2)^{n_i}\cdots(\lambda-\lambda_t)^{l_i},(i=1,2,\cdots,r)$$

则所有的

$$(\lambda-\lambda_1)^{m_i},(\lambda-\lambda_2)^{n_i},\cdots,(\lambda-\lambda_t)^{l_i},(i=1,2,\cdots,r),$$

除去幂指数为0的外，就恰好构成了 $B(\lambda)$ 的初等因子组。

证明 先证所有的 $(\lambda-\lambda_1)^{m_i},(i=1,2,\cdots,r)$，除指数为 0 的外，都是$B(\lambda)$的初等因子。设

$$D_1(\lambda),D_2(\lambda),\cdots,D_r(\lambda);\ d_1(\lambda),\ d_2(\lambda),\cdots,d_r(\lambda)$$

分别是 $B(\lambda)$ 的行列式因子与不变因子。再不妨设诸整数 m_i 是按大小顺序排好的：

$$0\leqslant m_1\leqslant m_2\leqslant\cdots\leqslant m_r.$$

这因为若不是这样，可对 $B(\lambda)$ 施行适当的初等变换（互换某两行，再互换这两列）就能使所得的 λ—矩阵（仍为对角形且与$B(\lambda)$相抵）对角线上的各多项式的因子 $(\lambda-\lambda_1)^{m_i}$ 的指数诸 m_i 满足我们的要求。这样假设好以后，$B(\lambda)$的任一k 阶子式就都能被$(\lambda-\lambda_1)^{m_1+m_2+\cdots+m_k}$整除，从而$(\lambda-\lambda_1)^m$ 便是 $D_k(\lambda)$ 的因子，其中 $m=m_1+m_2+\cdots+m_k$。但$(\lambda-\lambda_1)^{m+1}$ 就不能整除$D_k(\lambda)$，其原因是左上角那个 k阶子式就只含有 $(\lambda-\lambda_1)$ 的 m 次幂。因此，$D_k(\lambda)=(\lambda-\lambda_1)^m g_k(\lambda)$，其中 $g_k(\lambda)$ 不再含有因子 $\lambda-\lambda_1$，亦即

$$D_k(\lambda)=(\lambda-\lambda_1)^{m_1+m_2+\cdots+m_k}g_k(\lambda),\lambda-\lambda_1\nmid g_k(\lambda).$$

同理可知，

$$D_{k-1}(\lambda)=(\lambda-\lambda_1)^{m_1+m_2+\cdots+m_{k-1}}g_{k-1}(\lambda),\ \ \lambda-\lambda_1\nmid g_{k-1}(\lambda).$$

而

$$d_k(\lambda)=\frac{D_k(\lambda)}{D_{k-1}(\lambda)}=(\lambda-\lambda_1)^{m_k}g_0(\lambda),\ \ \lambda-\lambda_1\nmid g_0(\lambda),$$

即不变因子 $d_k(\lambda)$ 中恰含$(\lambda-\lambda_1)$的 m_k 次幂，故 $(\lambda-\lambda_1)^{m_k}$ 当 $m_k\neq 0$ 时为 $B(\lambda)$ 的一个初等因子，$k=1,2,\cdots,r$.

同理可证定理中其他所有因子（除指数为 0 者外）都是 $B(\lambda)$ 的初等因子.

次证除了这些因子外，$B(\lambda)$ 不再有其他的初等因子. 从上面的证明过程已经看出：如果$(\lambda-\lambda_0)^{m_0}$ 是含于某个不变因子中的初等因子，而 λ_0 又是$\lambda_1,\cdots,\lambda_t$ 中之一，则这个因子就已经是定理中的那些因子之一了. 所以，只要证明 $B(\lambda)$ 的初等因子组中不含有$(\lambda-\lambda_0)^{m_0}$，其中 $\lambda_0\neq\lambda_i$ $(i=1,2,\cdots,t)$，这样的初等因子就够了. 这是明显的事实，因为 $(\lambda-\lambda_0)$在 $\lambda_0\neq\lambda_i$ $(i=1,2,\cdots,t)$ 时，就不是诸 $h_i(\lambda)$的任何一个因子，从而不是 $B(\lambda)$ 中任何一个行列式因子的因子，自然更不是 $B(\lambda)$ 中任何一个不变因子的因子了. 于是定理得证.

这个定理告诉我们，一个对角形 λ—矩阵的初等因子组恰是对角线上诸多项式（在复数域中）的标准分解式中所有非零次幂的因子的全体. 由于相抵的两个 λ—矩阵初等因子组和秩相同，所以求一般的 λ—矩阵 $A(\lambda)$ 的初等因子组，可以对$A(\lambda)$ 施行初等变换化为对角形 λ—矩阵，对角线上非零多项式的个数就是 $A(\lambda)$的秩，而对角线上诸多项式（在复数域中）的标准分解式中所有非零次幂的因子恰好是$A(\lambda)$的初等因子组. 例如，求 λ—矩阵

$$A(\lambda)=\begin{pmatrix}0 & \lambda^2(\lambda-2)^3 & 0 & 0 & 0\\ \lambda(\lambda-2)^2 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & \lambda-2\\ 0 & 0 & \lambda(\lambda+1) & 0 & 0\\ 0 & 0 & 0 & \lambda-2 & 0\end{pmatrix}$$

的初等因子组，不变因子。

解 对 $A(\lambda)$ 用初等变换化为对角形 λ—矩阵：即交换 $A(\lambda)$ 的一二两行，三四两行，得

$$A(\lambda)\to\begin{pmatrix}\lambda(\lambda-2)^2 & 0 & 0 & 0 & 0\\ 0 & \lambda^3(\lambda-2)^3 & 0 & 0 & 0\\ 0 & 0 & \lambda(\lambda+1) & 0 & 0\\ 0 & 0 & 0 & 0 & \lambda-2\\ 0 & 0 & 0 & \lambda-2 & 0\end{pmatrix}.$$

再交换四五两行，得

$$A(\lambda)\to\begin{pmatrix}\lambda(\lambda-2)^2 & & & & \\ & \lambda^3(\lambda-2)^3 & & & \\ & & \lambda(\lambda+1) & & \\ & & & \lambda-2 & \\ & & & & \lambda-2\end{pmatrix}.$$

于是 $A(\lambda)$ 的秩为 5；初等因子组为 $\lambda,(\lambda-2)^2$，$\lambda^3,(\lambda-2)^3$，λ，$\lambda+1$，$\lambda-2,\lambda-2$；而不变因子为

$$d_5(\lambda)=\lambda^3(\lambda-2)^3(\lambda+1)$$

$$d_4(\lambda)=\lambda(\lambda-2)^2$$

$$d_3(\lambda)=\lambda(\lambda-2)$$

$$d_2(\lambda)=\lambda-2$$

$$d_1(\lambda)=1.$$

下面我们来解决第二个问题，即由A的不变因子，初等因子组所决定的有理型矩阵、若当型矩阵都与A是相似的。解决这个问题可分为两步：第一步证明A的有理型矩阵与A有相同的不变因子，A的若当型矩阵与A有相同的初等因子组。总括起来就是证明A的特征矩阵与有理型矩阵、若当型矩阵的特征矩阵相抵；第二步证明两个数字矩阵相似的充分必要条件是其特征矩阵相抵。

先来解决第一步的问题。

命题 3 A与其有理型矩阵有相同的不变因子。

证明 设A的不变因子为

$$1,\cdots,1,d_1,\cdots,d_s(\lambda),$$

其中每个 $d_i(\lambda)$ 都是非零次的，且每个

$$d_i(\lambda)=\lambda^{n_i}+a_{i1}\lambda^{n_i-1}+\cdots+a_{in_{i-1}}\lambda+a_{in_i}$$

的伴侣矩阵为

$$N_i=\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_{in_i} \\ 1 & 0 & 0 & \cdots & 0 & -a_{in_{i-1}} \\ 0 & 1 & 0 & \cdots & 0 & -a_{in_{i-2}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{i2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{i1} \end{pmatrix}.$$

其有理型矩阵为

$$B=\begin{pmatrix} N_1 & & \\ & \ddots & \\ & & N_s \end{pmatrix}.$$

现在求

$$\lambda I_n-B=\begin{pmatrix} \lambda I_{n_1}-N_1 & & \\ & \ddots & \\ & & \lambda I_{n_s}-N_s \end{pmatrix}$$

的不变因子，其中 I_{ni} 表示 n_i 阶单位阵。

因为

$$\lambda I_{ni}-N_i=\begin{pmatrix} \lambda & 0 & 0 & \cdots & 0 & a_{in_i} \\ -1 & \lambda & 0 & \cdots & 0 & a_{in_{i-1}} \\ 0 & -1 & \lambda & \cdots & 0 & a_{in_{i-2}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda & a_{i2} \\ 0 & 0 & 0 & \cdots & -1 & \lambda+a_{i1} \end{pmatrix}$$

的各阶行列式因子为

$D_1(\lambda)=1,\cdots,D_{n_{i-1}}(\lambda)=1,\quad D_{ni}(\lambda)=\lambda^{ni}+a_{i1}\lambda^{n_i-1}+\cdots+a_{in_{i-1}}\lambda+a_{ni}$，于是 N_i 的不变因子为 $1,\cdots,1,d_i(\lambda)$。故

$$\lambda I_{ni}-N_i\to\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d_i(\lambda) \end{pmatrix},$$

所以

$$\lambda I_n - B \to \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & d_1(\lambda) & & & \\ & & & & \ddots & & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1 \\ & & & & & & & & d_s(\lambda) \end{pmatrix}$$

$$\to \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & d_1(\lambda) & & \\ & & & & \ddots & \\ & & & & & d_s(\lambda) \end{pmatrix}$$

故 $\lambda I_n - B$ 即 B 的不变因子也为

$$1, \cdots, 1, d_1(\lambda), \cdots, d_s(\lambda).$$

这就证明了本命题。

命题 4　A 与其若当型矩阵有相同的初等因子组。

证明　设 A 的初等因子组为

$$(\lambda - \rho_1)^{m_1}, (\lambda - \rho_2)^{m_2}, \cdots, (\lambda - \rho_t)^{m_t}.$$

而每个初等因子 $(\lambda - \rho_i)^{m_i}$ 的若当块为 m_i 阶阵：

$$J_i = \begin{pmatrix} \rho_i & 1 & & \\ & \rho_i & 1 & \\ & & \ddots & \ddots \\ & & & 1 \\ & & & \rho_i \end{pmatrix}.$$

则 A 的若当型阵为

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{pmatrix}.$$

现在来求

$$\lambda I_n - J = \begin{pmatrix} \lambda I_{m_1} - J_1 & & & \\ & \lambda I_{m_2} - J_2 & & \\ & & \ddots & \\ & & & \lambda I_{m_t} - J_t \end{pmatrix}$$

的初等因子组。因为

$$\lambda I_{m_i}-J_i=\begin{pmatrix}\lambda-\rho_i & -1 & & \\ & \lambda-\rho_i & -1 & \\ & & \ddots & \ddots \\ & & & -1 \\ & & & \lambda-\rho_i\end{pmatrix}$$

的各阶行列式因子为

$$D_1(\lambda)=1,\cdots,\ D_{m_{i-1}}(\lambda)=1,\ D_{m_i}(\lambda)=(\lambda-\rho_i)^{m_i},$$

所以 J_i 的不变因子为

$$1,\cdots,1,\ d_i(\lambda)=(\lambda-\rho_i)^{m_i}.$$

故

$$\lambda I_{mi}-J_i\to\begin{pmatrix}1 & & \\ & \ddots & \\ & & 1 \\ & & & (\lambda-\rho_i)^{mi}\end{pmatrix}.$$

因此

$$\lambda I_n-J\to\begin{pmatrix}1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & (\lambda-\rho_1)^{m_1} & & & & & & \\ & & & & 1 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 1 & & & \\ & & & & & & & (\lambda-\rho_2)^{m_2} & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & 1 \\ & & & & & & & & & & \ddots \\ & & & & & & & & & & & 1 \\ & & & & & & & & & & & & (\lambda-\rho_t)^{m_t}\end{pmatrix}$$

由定理 4 可知 λI_n-J，即 J 的初等因子组恰为

$$(\lambda-\rho_1)^{m_1},(\lambda-\rho_2)^{m_2},\cdots,(\lambda-\rho_t)^{m_t}.$$

证完。

再来解决第二步的问题。为此，先给出两个命题做准备。

命题 5 设A,B 为 n 阶数字矩阵。如果存在数字矩阵 P_0,Q_0使

$$\lambda I-A=P_0(\lambda I-B)Q_0$$

则A与B相似。

证明　由

$$\lambda I - A = P_0(\lambda I - B)Q_0 = \lambda P_0 Q_0 - P_0 B Q_0$$

于是比较关于 λ 的矩阵多项式的系数，则有

$$P_0 Q_0 = I,\quad A = P_0 B Q_0$$

即 $P_0 = Q_0^{-1}$，$A = Q_0^{-1} B Q_0$，所以A与B相似，证完。

命题6　设 $U(\lambda)$为任一 λ—矩阵，A为非零数字矩阵．那么存在 λ—矩阵 $Q(\lambda)$，数字矩阵R，使

$$U(\lambda) = (\lambda I - A)Q(\lambda) + R,$$

类似地，有$\overline{Q}(\lambda)$，$\overline{R}$，使

$$U(\lambda) = \overline{Q}(\lambda)(\lambda I - A) + \overline{R}.$$

证明　这里和多项式的综合除法一样，只是由于矩阵乘法不满足交换律，从而除得的商与余式有左、右之别．我们只证前一个等式，因为后者与前者完全类似．

把 λ—矩阵 $U(\lambda)$ 表达成以矩阵为系数的多项式的形式：

$$U(\lambda) = D_0\lambda^m + D_1\lambda^{m-1} + \cdots + D_{m-1}\lambda + D_m$$

其中 $D_0, D_1, \cdots, D_{m-1}, D_m$都是数字矩阵，且$D_0 \neq 0$．

如果 $m = 0$，取 $Q(\lambda) = 0$，$R = D_0$

现在设 $m > 0$，用待定系数法．令

$$Q(\lambda) = Q_0\lambda^{m-1} + Q_1\lambda^{m-2} + \cdots + Q_{m-2}\lambda + Q_{m-1},$$

于是

$$(\lambda I - A)Q(\lambda) = Q_0\lambda^m + (Q_1 - AQ_0)\lambda^{m-1} + \cdots + (Q_{m-1} - AQ_{m-2})\lambda - AQ_{m-1}.$$

取

$$\begin{aligned}
Q_0 &= D_0\\
Q_1 &= D_1 + AQ_0\\
Q_2 &= D_2 + AQ_1\\
&\ \vdots\\
Q_{m-1} &= D_{m-1} + AQ_{m-2}\\
R &= D_m + AQ_{m-1}.
\end{aligned}$$

便有

$$U(\lambda)=(\lambda I-A)Q(\lambda)+R$$

成立。证完.

有了命题5、6的准备，我们便可证明以下重要的

定理5 两个数字方阵A与B相似的充分必要条件是 $\lambda I-A$ 与 $\lambda I-B$ 相抵。

证明 先证必要性，由 $A\sim B$，有可逆矩阵P，使$B=P^{-1}AP$。而

$$\lambda I-B=\lambda I-P^{-1}AP=P^{-1}(\lambda I-A)P.$$

据本节推论2知，$\lambda I-A\rightarrow\lambda I-B$.

再来证充分性. 设 $\lambda I-A$ 与 $\lambda I-B$相抵，于是有可逆 λ—矩阵 $U(\lambda)$，$V(\lambda)$，使

$$\lambda I-A=U(\lambda)(\lambda I-B)V(\lambda).$$

据命题6，可把 $U(\lambda)$、$V(\lambda)$写成以下形式：

$$U(\lambda)=(\lambda I-A)Q(\lambda)+R_1 \qquad (2)$$

$$V(\lambda)=P(\lambda)(\lambda I-A)+R_2. \qquad (3)$$

把（3）代入

$$U^{-1}(\lambda)(\lambda I-A)=(\lambda I-B)V(\lambda)$$

经整理，得

$$[U^{-1}(\lambda)-(\lambda I-B)P(\lambda)](\lambda I-A)=(\lambda I-B)R_2.$$

上式右端看成一个矩阵多项式，其次数≤1. 因此左端的

$$U^{-1}(\lambda)-(\lambda I-B)P(\lambda)$$

必为数字矩阵，记作 $T=U^{-1}(\lambda)-(\lambda I-B)P(\lambda)$，于是有

$$U(\lambda)T+U(\lambda)(\lambda I-B)P(\lambda)=I \qquad (4)$$

$$T(\lambda I-A)=(\lambda I-B)R_2. \qquad (5)$$

现在我们证明T是可逆的. 由（4）式知，

$$\begin{aligned}I&=U(\lambda)T+U(\lambda)(\lambda I-B)P(\lambda)\\&=U(\lambda)T+(\lambda I-A)V^{-1}(\lambda)P(\lambda)\\&=[(\lambda I-A)Q(\lambda)+R_1]T+(\lambda I-A)V^{-1}(\lambda)P(\lambda)\end{aligned}$$

$$= R_1 T + (\lambda I - A)[Q(\lambda)T + V^{-1}(\lambda)P(\lambda)].$$

考虑上式两端作为矩阵多项式的次数，由于 R_1T，I都是数字矩阵，于是必有

$$Q(\lambda)T + V^{-1}(\lambda)P(\lambda) = 0.$$

故得

$$R_1T = I，即T是可逆的。$$

从而由（5）式，可得

$$\lambda I - A = T^{-1}(\lambda I - B)R_1$$

由命题5，得A与B相似，证完。

由定理5，立刻得到以下结论：

两个 n 阶数字矩阵A与B相似的充分必要条件是A与B的不变因子相同；

两个 n 阶数字矩阵A与B相似的充分必要条件是A与B的初等因子组相同；

两个 n 阶数字矩阵A与B相似的充分必要条件是A与B的行列式因子相同。

这表明：不变因子、初等因子组、行列式因子都是相似分类的完备不变量。

由于数字矩阵A与它的有理型阵不变因子相同，A与它的若当型阵的初等因子组相同，所以，A与其有理型阵相似，A与其若当型矩阵相似。

到现在为止，我们提出的两个问题都得到了完满的解决。最后借助于若当标准形，再给出：一个n阶阵A相似于对角形阵的充分必要条件。

我们知道，任一n阶阵A都与一个若当型阵相似，而若当型阵对角线上诸块都是一阶块时，就是对角形阵，因此有

定理6　A与对角形阵相似的充分必要条件是A的初等因子都是一次的。

证明　先证必要性。设

$$A \sim B = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

由本节定理 5 知，

$$\lambda I_n - A \to \lambda I_n - B = \begin{pmatrix} \lambda-\lambda_1 & & & \\ & \lambda-\lambda_2 & & \\ & & \ddots & \\ & & & \lambda-\lambda_n \end{pmatrix}.$$

于是由本节定理 4 知，B 的初等因子恰好是

$$\lambda-\lambda_1, \lambda-\lambda_2, \cdots, \lambda-\lambda_n,$$

都是一次的. 故A的初等因子都是一次的.

再证充分性. 由于A的初等因子都是一次的，故每个初等因子决定的若当块都是一阶块，从而A的若当型阵为对角形阵，也就是A与对角形阵相似，证完.

与对角形阵相似的充分必要条件还可以通过最小多项式来得到. 为此，我们先证

定理 7 n **阶阵A的最后一个不变因子就是 A的最小多项式.**

证明 设A的若当型阵

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_t \end{pmatrix},$$

其中 J_i 是由初等因子 $(\lambda-\rho_i)^{m_i}$ 决定的若当块. 显然A与J有相同的最小多项式，有相同的不变因子. 按定理 7 的要求，只须证J的最小多项式就是J的最后一个不变因子即可.

事实上，我们知道J的最小多项式是诸块 $J_1, J_2, \cdots, J_t$ 的最小多项式的最小公倍式，而每个块 J_i 的最小多项式恰为 $(\lambda-\rho_i)^{m_i}$. 这因为

$$(J_i-\rho_i I_{m_i})^{m_i} = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & 0 \end{pmatrix}^{m_i} = 0,$$

并且当 $k<m_i$ 时，

$$(J_i-\rho_i I_{m_i})^k=\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & & 1 \\ & & & & 0 \end{pmatrix}^k \neq 0.$$

这就是说，J 的最小多项式恰好等于

$$(\lambda-\rho_1)^{m_1},\ (\lambda-\rho_2)^{m_2},\cdots,(\lambda-\rho_t)^{m_t}$$

的最小公倍式.

另外，J 的初等因子组恰为

$$(\lambda-\rho_1)^{m_1},(\lambda-\rho_2)^{m_2},\cdots,(\lambda-\rho_t)^{m_t}.$$

而 J 的最后一个不变因子的标准分解式中包含了初等因子组中所有根互异的初等因子，根相同的取方指数最高的那一个. 于是 J 的最后一个不变因子恰好等于

$$(\lambda-\rho_1)^{m_1},(\lambda-\rho_2)^{m_2},\cdots,(\lambda-\rho_t)^{m_t}$$

的最小公倍式. 故本定理得证.

定理 8　n 阶阵 A 相似于对角形阵的充分必要条件是 A 的最小多项式无重根.

证明　由本节定理 6 知，A 与对角形阵相似的充分必要条件是 A 的初等因子都是一次的. A 的初等因子都是一次的充分必要条件是 A 的最后一个不变因子无重因式，也就是 A 的最后一个不变因子无重根. 由本节定理 7 知，A 的最小多项式无重根，定理得证.

定理 8 告诉我们：判断一个阵 A 能否与对角形阵相似，只须看 A 的最小多项式是否无重根.

练　习　八

1. 判断下列 $A(\lambda)$ 与 $B(\lambda)$ 是否相抵：

1）
$$A(\lambda)=\begin{pmatrix} 8\lambda+1 & \lambda & 4\lambda-1 \\ 1-\lambda^2 & \lambda-1 & \lambda-\lambda^2 \\ \lambda^2+\lambda+2 & \lambda & \lambda^2+2\lambda \end{pmatrix},$$

$$B(\lambda)=\begin{pmatrix}\lambda+1 & \lambda-2 & \lambda^2-2\lambda\\ 2\lambda & 2\lambda-3 & \lambda^2-2\lambda\\ -2 & 1 & 1\end{pmatrix}$$

2）

$$A(\lambda)=\begin{pmatrix}\lambda-a & 0 & 0 & -1 & 0 & 0\\ 0 & \lambda-a & 0 & 0 & -1 & 0\\ 0 & 0 & \lambda-a & 0 & 0 & -1\\ \beta^2 & 1 & 0 & \lambda-a & 0 & 0\\ 0 & \beta^2 & 1 & 0 & \lambda-a & 0\\ 0 & 0 & \beta^2 & 0 & 0 & \lambda-a\end{pmatrix},$$

$$B(\lambda)=\begin{pmatrix}-1 & 0 & 0 & 0 & 0 & 0\\ 0 & -1 & 0 & 0 & 0 & 0\\ 0 & 0 & -1 & 0 & 0 & 0\\ 0 & 0 & 0 & (\lambda-a)^2+\beta^2 & 1 & 0\\ 0 & 0 & 0 & 0 & (\lambda-a)^2+\beta^2 & 1\\ 0 & 0 & 0 & 0 & 0 & (\lambda-a)^2+\beta^2\end{pmatrix}.$$

2．求

$$A(\lambda)=\begin{pmatrix}\lambda-a & c_1 & & & \\ & \lambda-a & c_2 & & \\ & & \ddots & \ddots & \\ & & & & c_{n-1}\\ & & & & \lambda-a\end{pmatrix}$$

的不变因子和初等因子组，其中 c_i 为非零常数，$(i=1,2,\cdots,n-1)$。

3．设 $(f(\lambda),g(\lambda))=1$，证明下面各 λ—矩阵是相抵的：

$$\begin{pmatrix}f(\lambda) & \\ & g(\lambda)\end{pmatrix},\ \begin{pmatrix}g(\lambda) & \\ & f(\lambda)\end{pmatrix},\ \begin{pmatrix}1 & \\ & f(\lambda)g(\lambda)\end{pmatrix}.$$

4．设 $A(\lambda)$ 为5阶 λ-矩阵，秩为4，初等因子组为 $\lambda,\lambda^2,\lambda^2,\lambda-1,\lambda-1,\lambda+1,(\lambda+1)^3$。试求 $A(\lambda)$ 的不变因子，并写出 $A(\lambda)$ 在相抵之下的标准形。

5．证明

1）n阶 λ—矩阵 $A(\lambda)$ 是可逆的充分必要条件是 $A(\lambda)$仅经行的初等变换就能化为 I_n；

2）n阶λ—矩阵 $A(\lambda)$ 是可逆的充分必要条件是 $A(\lambda)$仅经列的初等变换就能化为 I_n。

6．判断下列各阵是否可逆，并将可逆阵表为初等阵之积：

1) $\begin{pmatrix} \lambda^2 & \lambda & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, 2) $\begin{pmatrix} 1 & \lambda & \lambda^2 \\ 0 & 2 & \lambda \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$, 3) $\begin{pmatrix} 5\lambda+1 & 25\lambda \\ \lambda & 5\lambda-1 \end{pmatrix}$,

4) $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$.

7. 秩为 n 的 n 阶 λ—矩阵称为满秩的. 试问满秩的 n 阶 λ—矩阵是否一定可逆？下列矩阵哪个是满秩的，哪个是可逆的？

$$A(\lambda)=\begin{pmatrix} 1 & 0 & 1 \\ 0 & \lambda-1 & \lambda \\ \lambda & 1 & \lambda^2 \end{pmatrix},\quad B(\lambda)=\begin{pmatrix} 1 & \lambda & 0 \\ 2 & \lambda & 1 \\ \lambda^2+1 & 2 & \lambda+3 \end{pmatrix}.$$

8. 证明：两个相抵的 λ—矩阵的行列式只相差一个非零的常数因子.

9. 证明下列方阵

$$A_1=\begin{pmatrix} r_1 & & \\ & r_1 & \\ & & r_1 \end{pmatrix},\quad A_2=\begin{pmatrix} r_1 & & \\ & r_1 & 1 \\ & & r_1 \end{pmatrix},\quad A_3=\begin{pmatrix} r_1 & 1 & \\ & r_1 & 1 \\ & & r_1 \end{pmatrix}$$

任何两个都不相似.

10. 判断下列两个矩阵是否相似：

$$A=\begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix},\quad B=\begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix}.$$

11. 证明：A 与其转置阵 A' 相似.

12. 设 n 阶阵 A，$d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)$ 为 A 的不变因子. 证明

$$|\lambda I_n - A| = d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda).$$

13. 在复数域中，n 阶方阵 A 的若当型阵中的若当块可否换为

$$\begin{bmatrix} \rho & a & & \\ & \rho & a & \\ & & \ddots & \ddots \\ & & & a \\ & & & \rho \end{bmatrix},\quad (a\neq 0)$$

形式，为什么？

习 题 十

1．证明：

1）相似矩阵或者同时可逆，或者同时都不可逆；

2）如果$A\sim B$，且A与B都可逆，则$A^{-1}\sim B^{-1}$。

2．证明：

1）与幂零阵相似的阵均为幂零阵；

2）与幺幂阵相似的阵均为幺幂阵；

3）与幂等阵相似的阵均为幂等阵。

3．设 $A_1\sim B_1$，$A_2\sim B_2,\cdots,A_s\sim B_s$。证明

$$\begin{pmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s\end{pmatrix}\sim\begin{pmatrix}B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_s\end{pmatrix}.$$

4．证明：n 阶矩阵A的行列式等于它的特征多项式的常数项的 $(-1)^n$ 倍，进而等于其特征根之积。

5．设 n 阶方阵A的任一行中 n 个元素之和都是 λ。证明，λ是A的特征根，$(1,\cdots,1)$是A的属于特征根 λ 的特征向量。

6．证明：反对称实矩阵的特征根是零或是纯虚数。

7．证明：在复数域C中，若 $AB=BA$，则A与B有公共的特征向量。

8．证明：n 阶方阵A的特征根必是A的最小多项式的根。

9．设n阶方阵A有n个不同的特征根。证明与A可换的n阶方阵B必与对角形阵相似。

10．设$e_1,e_2,\cdots,e_n$；$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 均为 n 维向量空间的标准正交基底，且

$$(e_1e_2\cdots e_n)=(\varepsilon_1\varepsilon_2\cdots\varepsilon_n)\begin{pmatrix}a_{11}a_{12}\cdots a_{1n}\\ a_{21}a_{22}\cdots a_{2n}\\ \cdots\cdots\cdots\cdots\\ a_{n1}a_{n2}\cdots a_{nn}\end{pmatrix}.$$

证明，

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

为正交矩阵。

11．设A、B都是正交矩阵．若 $|A|=-|B|$，证明$|A+B|=0$。

12．设A、B都是正交矩阵．证明：

1）方程 $|\lambda A-B|=0$的根均为正的；

2）$A=B$的充分必要条件是$|\lambda A-B|=0$的所有根均为1。

13．二次型$f(x_1,\cdots,x_n)$的表示矩阵为A，λ是A的特征根．证明，存在非零的n维向量$(\zeta_1,\zeta_2,\cdots,\zeta_n)$，使

$$f(\zeta_1,\zeta_2,\cdots,\zeta_n)=\lambda(\zeta_1^2+\zeta_2^2+\cdots+\zeta_n^2)。$$

14．设S为实反对称阵．证明：

1）$I-S$和$I+S$都是可逆的；

2）$A=(I-S)(I+S)^{-1}$为正交矩阵．

15.设A为实对称阵，且$A^2=A$。证明，存在正交矩阵T，使

$$T^{-1}AT=\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}.$$

16．设A为三阶正交矩阵，且$\det A=1$．证明，存在一个数$t(1\leqslant t\leqslant 3)$，使

$$A^3-tA^2-tA-I=0。$$

17．设

$$A=\begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix},\quad B=\begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix}$$

求与 A^{100}、B^{100} 相似的阵。

18. 设A为 n 阶方阵。证明：A与对角形阵相似的充分必要条件是存在一个次$\geqslant 1$的无重根的多项式$f(\lambda)$，使 $f(A)=0$。

19. 证明：

1）非零的幂零矩阵不能与对角形阵相似；

2）幂零阵A，有 $|A+I|=1$。

20. 证明：幺幂方阵A与对角形阵相似。

21. 求出三阶对合阵的一切可能的若当标准形。

22. 求出三阶方阵A的若当标准形的一切可能形式。

23. 设 n 阶方阵A的最小多项式为$\varphi(\lambda)$，A的 $n-1$ 阶行列式因子 $D_{n-1}(\lambda)$ 为 $n-1$ 次的多项式。证明：$D_{n-1}(\lambda)=[\varphi(\lambda)]^{n-1}$。

24. 设A为 n 阶方阵。证明：A与对角形阵相似的充分必要条件是对A的任一特征根 λ_i，均有 $(\lambda_i I_n-A)^2$ 与 $(\lambda_i I_n-A)$ 秩数相同。

25. 设A为 n 阶方阵。证明：A与对角形阵相似的充分必要条件是对 A的任一特征根 λ_i，方程组

$$(\lambda_i I_n-A)^2\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}\text{与}\quad(\lambda_i I_n-A)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$$

是同解的。

26. 设二阶实方阵A的行列式为负数。证明，存在二阶满秩实阵P，使$P^{-1}AP$ 为对角形。

27. 若实数域上的两个 n 阶方阵A，B在复数域上相似，证明A与B必在实数域上相似。

第十一章　线性空间

§1　线性空间的定义与简单性质

在第六章为了研究线性方程组解的结构给出了n维向量空间$F^{(n)}$的概念．做成空间的元素是n数组（写成$1\times n$矩阵的形式或写成$n\times 1$矩阵的形式皆可），对空间的元素n数组可以做两种运算：加法运算与倍数运算，并且满足熟知的八个条件：

1）加法交换律

$\alpha+\beta=\beta+\alpha,$

2）加法结合律

$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma),$

3）有零向量0，对任意的向量α使

$\alpha+0=\alpha=0+\alpha,$

4）对每一个向量α，存在一个向量α'使

$\alpha+\alpha'=0=\alpha'+\alpha,$

5）$k(\alpha+\beta)=k\alpha+k\beta,$

6）$(k+l)\alpha=k\alpha+l\alpha,$

7）$(kl)\alpha=k(l\alpha),$

8）$1\cdot\alpha=\alpha.$

其中α，β，γ为$F^{(n)}$中任意向量，k，l为F中任意的数．

综合起来，可以概括地说，数域F上的向量空间$F^{(n)}$由三个要素组成：

1．一组可运算的元素n数组；

2．两个运算方法：加法运算与倍数运算；

3．八个条件：1）～8）．

适合以上三方面要求的除了由n数组构成的向量空间之外，大家比较熟悉的，数域F上一元多项式集合$F[x]$和F上$m\times n$矩阵集合$M_{m\times n}(F)$也都全面满足上述三方面的要求．实际上，满足以上三方面要求的是一类相当广泛的数学对象．为了对这类对象用统一的方法加以研究，以便形成一般的理论，我们自然地引出线性空间的概念．

定义　假设

1．F是任意数域，V是任意非空集合．

2．V有一个运算，叫加法；F，V到V有一个运算，叫倍数乘法．

3．这两个运算满足下列八个条件：

1）$\alpha+\beta=\beta+\alpha$,

2）$(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$,

3）V中有元素θ，$\forall\alpha\in V$使

$\alpha+\theta=\alpha=\theta+\alpha$.

4）对每一$\alpha\in V$，有$\alpha'\in V$，使

$\alpha+\alpha'=\theta=\alpha'+\alpha$.

5）$k(\alpha+\beta)=k\alpha+k\beta$,

6）$(k+l)\alpha=k\alpha+l\alpha$,

7）$(kl)\alpha=k(l\alpha)$,

8）$1\cdot\alpha=\alpha$.

其中α，β，γ为V中任意元素，k，l为F中任意数．则称V做成F上的线性空间（也说V是F上的线性空间），V中的元素叫做向量．以下常用$V(F)$表示V是F上的线性空间．

例1　由F上n数组构成的向量空间$F^{(n)}$是F上的线性空间．

例2　$F[x]$：$f(x)=a_0+a_1x+\cdots+a_nx^n$，$a_i\in F$，$n$为任意非负整数，按多项式的加法，数与多项式相乘的倍数运算，$F[x]$是

F 上的线性空间，$F[x]$ 的每一多项式都叫做向量。

例 3　$M_{m\times n}(F)$ 按矩阵加法，数与矩阵相乘的倍数运算做成 F 上的线性空间. 这时每一 $m\times n$ 矩阵都叫线性空间 $M_{m\times n}(F)$ 的向量。

例 4　$C[a, b]$ 表示定义在闭区间 $[a, b]$ 上一切连续函数的集合. 按函数的加法，数与函数相乘的倍数运算是实数域 D 上的线性空间。

例 5　取 F 为实数域 D，V 为复数域 C。于是按复数的加法，实数与复数相乘的倍数运算，复数域 C 做成实数域 D 上的线性空间。

从线性空间的定义直接可以推出以下一些简单而基本的性质。

1° 有唯一的零向量. 具有性质

$$\xi+\theta=\xi=\theta+\xi,\quad \forall\xi\in V$$

的向量 θ 是唯一的。

事实上　假设 V 中向量 $\bar\theta$ 也有上述性质，即

$$\xi+\bar\theta=\xi=\bar\theta+\xi,\quad \forall\xi\in V$$

于是考察 $\theta+\bar\theta$，就同时有

$$\theta+\bar\theta=\bar\theta,\quad \theta+\bar\theta=\theta,$$

所以，$\bar\theta=\theta$。

把这个唯一的向量 θ 叫做 V 的零向量，简记作 0 。

2° 每一个向量都有唯一的负向量. 对向量 ξ 使

$$\xi+\xi'=0=\xi'+\xi$$

的向量 ξ' 是唯一的。

事实上，假设 V 中向量 ξ'' 也有上述性质，即

$$\xi+\xi''=0=\xi''+\xi$$

于是考察

$\xi'+\xi+\xi''$，就同时有

$$\xi'+\xi+\xi''=(\xi'+\xi)+\xi''=0+\xi''=\xi'',$$

$$\xi'+\xi+\xi''=\xi'+(\xi+\xi'')=\xi'+0=\xi'.$$

所以，$\xi''=\xi'$.

把这个唯一的向量ξ'叫做向量ξ的负向量，记作 $\xi'=-\xi$.

3° $0\cdot\alpha=0$，$k0=0$

$$k(-\alpha)=-(k\alpha)=(-k)\alpha.$$

先证明 $0\cdot\alpha=0$，$k0=0$. 我们有

$$0\cdot\alpha+0\cdot\alpha=(0+0)\alpha=0\cdot\alpha$$

两边加上$-(0\cdot\alpha)$ 即得

$$0\cdot\alpha=0$$

$$k0+k0=k(0+0)=k0,$$

两边加上$-(k0)$ 即得

$$k0=0.$$

其次证明 $k(-\alpha)=-(k\alpha)=(-k)\alpha$

$$k\alpha+k(-\alpha)=k(\alpha+(-\alpha))=k0=0,$$

按负向量的定义即得

$$k(-\alpha)=-(k\alpha)$$

$$(-k)\alpha+k\alpha=(-k+k)\alpha=0\alpha=0,$$

按负向量的定义即得

$$(-k)\alpha=-(k\alpha).$$

特别地，$k=1$时就有

$$(-1)\alpha=-\alpha.$$

4° 若 $k\alpha=0$，则 $k=0$ 或 $\alpha=0$

设 $k\alpha=0$. 如果 $k\neq0$，去证必有 $\alpha=0$.

在等式 $k\alpha=0$ 的两端同时乘以$\frac{1}{k}$，则有

$$\frac{1}{k}(k\alpha)=\frac{1}{k}0$$

$$\left(\frac{1}{k}\cdot k\right)\alpha=0$$

$$1\alpha=0$$

$\alpha=0$

5° 设$\alpha\neq0$。若 $k\neq h$，则 $k\alpha\neq h\alpha$。

如果 $k\alpha=h\alpha$，在两端同时加上 $(-h)\alpha$，则有

$k\alpha+(-h)\alpha=h\alpha+(-h)\alpha=(h+(-h))\alpha=0\alpha=0$。

即得

$$(k-h)\alpha=0.$$

因为 $k\neq h$，所以 $k-h\neq0$。同时 $\alpha\neq0$。 这是一个矛盾。因之应有 $k\alpha\neq h\alpha$。

6° **向量的减法．由于每一向量都有唯一的负向量，这使我们可以定义向量的减法如下：**

$$\alpha-\beta=\alpha+(-\beta).$$

值得注意的是，一般来说向量的减法既不满足交换律，也不满足结合律。但是，倍数运算对减法是满足分配律的。即

$$k(\alpha-\beta)=k\alpha-k\beta.$$

事实上

$$\begin{aligned}k(\alpha-\beta)&=k(\alpha+(-\beta))=k\alpha+k(-\beta)\\&=k\alpha+(-k\beta)=k\alpha-k\beta.\end{aligned}$$

练 习 一

1．有没有一个向量的线性空间？有没有两个及有限个向量的线性空间？为什么？

2．按通常多项式的加法和数与多项式的乘法，下面集合是否构成数域F上的线性空间？

1) V是数域F上次数等于n的多项式的集合。

2) $V=\{ax_1^2+bx_1x_2+cx_2^2 \mid a, b, c\in F\}$

3．平面π上位于第一象限的所有向量，关于向量的加法和数乘向量的倍数乘法是否构成线性空间？

4．按通常矩阵的加法和数乘矩阵的倍数乘法，下列集合是否

是实数域D上的线性空间？

1）实数域上n阶对称阵的全体所构成的集合。

2）实数域上n阶反对称阵的集合。

3）$V=\{(a_{ij})\mid a_{11}+a_{22}+\cdots+a_{nn}=0,\ a_{ij}\in D\}$

4）$GL_n(D)$ 是实数域D上n阶可逆矩阵的集合。

5）$SL_n(D)$ 是实数域D上 n 阶方阵，且行列式的值等于1的集合。

5．平面π上全体向量所构成的集合，对于通常的向量加法和如下定义的数与向量的倍数乘法

$$k\cdot\alpha=0$$

其中α是集合中任一向量，$k\in D$．问该集合是否是实数域D上的线性空间？

6．按数的加法和乘法，下列集合是否是实数域D上的线性空间？

1）Z：整数集。

2）Q：有理数集。

3）D：实数集。

7．举例说明线性空间关于运算的第八条不能由前七条推出来。

§2 子空间

在研究线性方程组解的结构问题时，向量空间的子空间的重要作用是十分明显的．对一般线性空间的理论来说，子空间也是相当重要的概念。

定义1　设V是F上的线性空间，W是V的一个非空子集．如果用V的加法与倍数运算W本身也做成F上的线性空间，则称W是V的一个子空间。

按定义判断一个子集W是否做成子空间似乎必须验证两个方

面：一是验证V的加法与倍数运算都是子集W的运算，二是逐条验证八个条件都成立。当然，必须切实验证V的加法与倍数运算都是W的运算。但是，逐条验证八个条件都成立是不必要的。因为我们有以下的事实。即

命题 **设W是线性空间V的非空子集。那么，W是V的子空间当且仅当 $\forall \alpha, \beta \in W$，$k \in F$都有 $\alpha+\beta \in W$，$k\alpha \in W$。**

证明 必要性是显然的。下面证明充分性。设$\forall \alpha, \beta \in W$，$k \in F$都有 $\alpha+\beta \in W$，$k\alpha \in W$，去证W是子空间。这实际上就是说，只要V的运算也是W的运算，八个条件自然都成立。这里需要证明的只是零向量在W里，和W中每个向量的负向量也在W里这两条，其余六条是不证自明的。

事实上，任取 $\alpha \in W$， 再取F中的数 $k=0$，那么 $k\alpha=0\alpha=0 \in W$。即零向量在W里。再取F中的数$k=-1$，那么$k\alpha=(-1)\alpha=-\alpha \in W$，即$W$中任一向量$\alpha$的负向量$-\alpha$也在$W$里。总之，$W$是$V$的子空间。

例 1 对任一线性空间V来说，$W_1=V$与$W_2=\{0\}$即V本身与只含一个零向量的子集，是V的两个明显的子空间。一般地，把只有一个向量的线性空间叫零空间；不同于V的子空间叫做V的真子空间。

例 2 $F^{(n)}$ 的子集

$$W=\{(x_1, \cdots, x_m, 0, \cdots, 0) \mid x_1, \cdots, x_m \in F\},$$

其中m为一取定的整数，则W是$F^{(n)}$的一个子空间。

例 3 $F[x]$ 的子集

$$F_m[x]=\{a_0+a_1x+\cdots+a_{m-1}x^{m-1} \mid a_i \in F\}$$

其中m为一取定的整数，则 $F_m[x]$ 是$F[x]$ 的真子空间。

例 4 在 $C[0, 1]$ 中，$F[x]$ 看做$C[0, 1]$的子集，是$C[0, 1]$ 的子空间。

例 5 在 $M_2(F)$ 中， $W_1=\{A \in M_2(F) \mid A'=A\}$， $W_2=\{A \in M_2(F) \mid A'=-A\}$。那么 W_1，W_2 都是 $M_2(F)$ 的子空间。

这里W_1就是一切二阶对称阵做成的子集，W_2就是一切二阶反对称阵做成的子集.

若 $A, B\in W_1$，即有 $A'=A$，$B'=B$. 于是

$$(A+B)'=A'+B'=A+B,\ (kA)'=kA'=kA,$$

所以，$A+B\in W_1$，$kA\in W_1$. 由前述命题，W_1是 $M_2(F)$ 的子空间. 完全类似地可证 W_2 是子空间，具体证明留给读者做为练习.

下面介绍构造子空间的一般方法，它对研究线性空间的结构也是有益的.

设V为任一线性空间. S为V的任一非空子集. 令

$$L(S)=\{k_1\xi_1+k_2\xi_2+\cdots+k_t\xi_t \mid \xi_i\in S,\ k_i\in F,\ t\in N\}$$

用话来说就是 $L(S)$ 是子集 S 中一切可能的有限个向量的倍数和构成的子集. 不难证明 $L(S)$ 做成V的子空间.

事实上 任取 $\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_r\alpha_r$，$\beta=b_1\beta_1+b_2\beta_2+\cdots+b_s\beta_s$. 此处 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_1, \cdots, \beta_s\in S$. $a_1, \cdots, a_r, b_1, \cdots, b_s\in F$.

于是

$$\alpha+\beta=a_1\alpha_1+\cdots+a_r\alpha_r+b_1\beta_1+\cdots+b_s\beta_s\in L(S),$$

$$k\alpha=ka_1\alpha_1+\cdots+ka_r\alpha_r\in L(S).$$

由前述命题，$L(S)$ 是一个子空间.

定义 2　$L(S)$ 叫做由子集S生成的子空间，S叫做 $L(S)$ 的一个生成子集，特别地，当S是有限子集时，即 $S=\{\xi_1,\xi_2,\cdots,\xi_t\}$. 则$L(S)=L(\xi_1, \xi_2, \cdots, \xi_t)$ 叫做有限生成子空间. 称$\xi_1, \xi_2, \cdots, \xi_t$ 是$L(\xi_1, \xi_2, \cdots, \xi_t)$ 的一组生成元或一个生成组.

显然，有限生成的空间比较简单，因为这种空间中每一向量都可表成一组确定的向量的倍数和.

如果 $V=L(\xi_1, \xi_2, \cdots, \xi_t)$ 则称线性空间V是有限生成的.

例 6　在$F^{(n)}$中，取

$\varepsilon_1=(1, 0, 0, \cdots, 0)$

$\varepsilon_2=(0, 1, 0, \cdots, 0)$

……

$\varepsilon_n=(0,\ 0,\ 0,\ \cdots,\ 1)$.

则有 $F^{(n)}=L(\varepsilon_1,\ \varepsilon_2,\cdots,\ \varepsilon_n)$。所以 $F^{(n)}$ 是有限生成的。$\varepsilon_1,\ \varepsilon_2,\cdots,\ \varepsilon_n$ 是它的一个生成组。

例 7　考虑 $F[x]$,

$$F_m[x]=\{a_0+a_1x+\cdots+a_{m-1}x^{m-1}\mid a_i\in F\}.$$

由于 $F_m[x]=L(1,\ x,\ \cdots,\ x^{m-1})$，所以 $F_m[x]$ 是有限生成的子空间。$1,\ x,\ \cdots,x^{m-1}$是 $F_m[x]$ 的一个生成组。

例 8　$V=M_2(F)$。$W_1=\{A\in M_2(F)\mid A'=A\}$。取

$$\varepsilon_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},\ \varepsilon_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},\ \varepsilon_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix},\ \varepsilon_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

则有

$$M_2(F)=L(\varepsilon_{11},\ \varepsilon_{12},\ \varepsilon_{21},\ \varepsilon_{22}),\ W_1=L(\varepsilon_{11},\ \varepsilon_{22},\ \varepsilon_{12}+\varepsilon_{21}).$$

所以W_1是有限生成的子空间，$M_2(F)$ 本身则是有限生成的线性空间。同理，反对称阵组成的子空间W_2（见例 5）也是有限生成的，且$W_2=L(\varepsilon_{12}-\varepsilon_{21})$。

⑲

例 9　并不是任何线性空间都是有限生成的，$F[x]$ 就不是有限生成的。

我们用反证法来说明这个事实。假设$F[x]$ 是有限生成的，那么有有限个非零向量 $f_1(x),\ f_2(x),\ \cdots,\ f_t(x)$ 是 $F[x]$ 的一组生成元，这样，对线性空间 $F[x]$中的任一向量 $g(x)$ 必有

11

$$g(x)=k_1f_1(x)+k_2f_2(x)+\cdots+k_tf_t(x),\ k_i\in F.$$

因为生成元都是多项式：$f_1(x),\ f_2(x),\ \cdots,\ f_t(x)$，所以它们都有一个确定的次数。令这 t 个次数中最大的是m。于是取 $g(x)=x^{m+1}$ 时，上式就是不可能成立的。因为这时左端的次数为 $m+1$，而右端的次数$\leqslant m$。这个矛盾表明线性空间 $F[x]$ 不可能是有限生成的。

例10　有限生成的线性空间的生成组一般来说不只一组。如

$$F^{(n)}=L(\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n)$$

$$= L(2\varepsilon_1,\ 2\varepsilon_2,\ \cdots,\ 2\varepsilon_n)$$
$$= L(\varepsilon_1,\ 2\varepsilon_2,\ \cdots,\ n\varepsilon_n).$$

例11　设 S 是线性空间 V 的非空子集，那么 S 是子空间当且仅当 $L(S)=S$。

事实上，如果子集 S 已是子空间．由于

$$L(S)=\{k_1\xi_1+k_2\xi_2+\cdots+k_s\xi_s \mid \xi_i\in S, k_i\in F\}$$

即得 $L(S)\subseteq S$，而 $S\subseteq L(S)$ 是明显的，所以 $L(S)=S$。

反之，如果 $L(S)=S$，那么对 $\forall\alpha,\ \beta\in S,\ k,\ h\in F$ 都有

$$k\alpha+h\beta\in L(S)$$

即 $k\alpha+h\beta\in S$。由前述命题，S 是子空间。

练　习　二

1．判断 $F^{(n)}$ 的下列子集哪些是 $F^{(n)}$ 的子空间？

1) $S_1=\left\{(a_1,\ a_2,\ \cdots,\ a_n)\ \middle|\ \sum_{i=1}^{n}a_i=0\right\}$

2) $S_2=\left\{(a_1,\ a_2,\ \cdots,\ a_n)\ \middle|\ \sum_{i=1}^{n}a_i=1\right\}$

2．D^+ 为正实数集，对任意 $\alpha,\ \beta\in D^+,\ k\in D$，按如下规定的加法与倍数乘法

$$\alpha\oplus\beta=\alpha\cdot\beta$$
$$k\cdot\alpha=\alpha^k$$

是实数域 D 上的线性空间，而 D 又是 D 上的线性空间，问 D^+ 是否是 D 的子空间？

3．实数集 D 既可看作有理数域 Q 上的线性空间，又可以看作实数域 D 上的线性空间，问是否可以将 D 在有理数域 Q 上的线性空间看成 D 在实数域上的线性空间的子空间？

4．一个含有 n 个未知数的非齐次线性方程组的通解是否是

$F^{(n)}$ 的子空间？

5．设A是某一确定的n阶方阵，$C(A)$ 表示一切和A可换的矩阵所构成的集合，证明，$C(A)$ 是 $M_n(F)$ 的子空间。

6．设A是 $M_n(F)$ 中一确定的n阶方阵

$$L_A=\{X\in M_n(F)\mid XA=0\}$$

证明：L_A是 $M_n(F)$ 的子空间。

§3 子空间的交与和 直和

对于任意给定的两个子空间 W_1，W_2 都可自然得出其它的子空间来。通常有两种方法，即由交 $W_1\cap W_2$ 生成一个确定的子空间。由并 $W_1\cup W_2$ 生成一个确定的子空间。这就是下面要讲的交空间与和空间。

由子空间做“和空间”是研究线性空间结构的重要方法。

设V为任意线性空间，W_1，W_2 是V的任意两个子空间。容易证明交集$W_1\cap W_2$ 也是一个子空间，即 $L(W_1\cap W_2)=W_1\cap W_2$。因为对任意的$\alpha$，$\beta\in W_1\cap W_2$，即 α，$\beta\in W_1$ 同时α，$\beta\in W_2$。于是

$$\alpha+\beta\in W_1 \quad 同时 \quad \alpha+\beta\in W_2$$

从而 $\alpha+\beta\in W_1\cap W_2$。完全类似地，对任意的$k\in F$，$\alpha\in W_1\cap W_2$都有 $k\alpha\in W_1\cap W_2$，所以，交集 $W_1\cap W_2$ 是一个子空间，叫做W_1与W_2的交空间，简称为W_1与W_2的交。

例 1 考虑$F^{(3)}$，$W_1=\{(x, y, 0)\mid x, y\in F\}$，$W_2=\{(0,y,z)\mid y, z\in F\}$。如果把$F^{(3)}$ 看做是几何空间，W_1相当于 xy——坐标面，W_2相当于yz——坐标面。于是W_1与W_2的交$W_1\cap W_2=\{(0, y, 0)\mid y\in F\}$。不难想象，这个交子空间就相当于 y——坐标轴。

例 2 若$AX=0$，$BX=0$ 的解空间为W_1，W_2，则$\begin{pmatrix}A\\B\end{pmatrix}X=0$的解空间为$W_1\cap W_2$。

例 3 考虑 $M_2(F)$，$N_1=\left\{\begin{pmatrix}a & 0\\ c & b\end{pmatrix}\middle| a, b, c\in F\right\}$．

$$N_2=\left\{\begin{pmatrix}a & c\\ 0 & b\end{pmatrix} a,\ b,\ c\in F\right\},N_3=\left\{\begin{pmatrix}a & 0\\ 0 & b\end{pmatrix} a,\ b\in F\right\}.$$

容易想到N_1，N_2，N_3都是 $M_2(F)$ 的子空间．于是

$$N_1\cap N_2=N_3,\ N_1\cap N_3=N_2\cap N_3=N_3.$$

两个子空间的并集一般的不做成子空间．如例 1 中的两个子空间W_1与W_2的并集$W_1\cup W_2$就不是子空间．因为取 $\alpha=(1,\ 1,\ 0)$，$\beta=(0,\ 1,\ 1)$ 于是 α，$\beta\in W_1\cup W_2$．但$\alpha+\beta=(1,\ 2,\ 1)$ 既不在W_1里，也不在W_2里，所以$\alpha+\beta$不在并集$W_1\cup W_2$里．因此 $W_1\cup W^2$ 不是一个子空间． 然而以并集 $W_1\cup W_2$ 为生成组总可以生成一个确定的子空间．即

$L(W_1\cup W_2)\ =\{k_1\alpha_1+\cdots+k_r\alpha_r+h_1\beta_1+\cdots+h_t\beta_t\ |\ \alpha_i\in W_1$，$\beta_j\in W_2,\ k_i,\ h_j\in \mathrm{F}\}$.

从而

$$L(W_1\cup W_2)=\{\xi+\eta\mid \xi\in W_1,\ \eta\in W_2\}.$$

即$L(W_1\cup W_2)$ 恰好是由W_1的向量与W_2 的向量随意相加得到的一切和向量组成的．因而称$L(W_1\cup W_2)$ 为W_1与W_2 **的和空间**，也简称为W_1与W_2的和．记作$L(W_1\cup W_2)=W_1+W_2$.

例 4 W_1，W_2如例 1，则有 $W_1+W_2=F^{(3)}$

例 5 $L(S)+L(T)=L(S\cup T)$

事实上 $L(S)\subseteq L(S\cup T)$，$L(T)\subseteq L(S\cup T)$，因为$L(S\cup T)$是子空间，所以$L(S)+L(T)\subseteq L(S\cup T)$．又由 $S\subseteq L(S),T\subseteq L(T)$，$L(S\cup T)\subseteq L(S)+L(T)$，$L(S\cup T)=L(S)\cup L(T)$.

例 6 考虑：$M_2(F)$，$S=\{A\in M_2(F)\mid A'=A\}$，$T=\{A\in M_2(F)\mid A'=-A\}$．证明：$S+T=M_2(F)$，$S\cap T=\{0\}$.

证明 $S+T=M_2(F)$ 是明显的，我们证明：$S\cap T=\{0\}$．即S与T的交只有一个零阵．任取 $A\in S\cap T$．从而 $A\in S$与$A\in T$．即得$A'=A$ 与 $A'=-A$，这就得出

$$A=-A,\ \text{即}\ A+A=2A=0,\ A=0.$$

即 $S\cap T=\{0\}$.

关于两个子空间的交与和有以下的简单性质．设 W_1，W_2，W_3，W 都是V的子空间．

1° 若$W\subseteq W_1$，$W\subseteq W_2$，则 $W\subseteq W_1\cap W_2$；若 $W_1\subseteq W$，$W_2\subseteq W$，则 $W_1+W_2\subseteq W$．

2° $W_1\cap W_2=W_1$当且仅当$W_1\subseteq W_2$．当且仅当 $W_1+W_2=W_2$

3° $W_1\cap W_2=W_2\cap W_1$，$W_1+W_2=W_2+W_1$；

$(W_1\cap W_2)\cap W_3=W_1\cap(W_2\cap W_3)$，

$(W_1+W_2)+W_3=W_1+(W_2+W_3)$．

4° 如果$W\subseteq W_1$，W_2，并且当 $W_3\subseteq W_1$，W_2 时必有$W_3\subseteq W$则$W=W_1\cap W_2$．

如果 W_1，$W_2\subseteq W$，并且当 W_1，$W_2\subseteq W_3$ 时必有$W\subseteq W_3$ 则 $W=W_1+W_2$．

前三条性质都是明显的，做为练习读者可以给出它们的证明．下面我们证明性质4°．

首先 由 $W\subseteq W_1$，W_2 则有 $W\subseteq W_1\cap W_2$，另外 $W_1\cap W_2\subseteq W_1,W_2$，从而$W_1\cap W_2\subseteq W$，所以$W_1\cap W_2=W$．

其次 由 W_1，$W_2\subseteq W$，则有 $W_1+W_2\subseteq W$，另外 $W_1,W_2\subseteq W_1+W_2$ 从而$W\subseteq W_1+W_2$，所以 $W_1+W_2=W$．

子空间的交与和很容易推广到任意有限个子空间的情形，设 W_1，W_2，…，W_r 是V的r个子空间，于是

$$W_1\cap W_2\cap\cdots\cap W_r=\bigcap_{i=1}^{r}W_i=\{\xi\mid\xi\in W_i;\ i=1,2,\cdots,r\},$$

容易证明$\bigcap\limits_{i=1}^{r}W_i$是$V$的一个子空间，叫做 W_1，W_2，…，W_r 的交（空间）．

$$W_1+W_2+\cdots+W_r=\sum_{i=1}^{r}W_i=\{\xi_1+\xi_2+\cdots+\xi_r\mid\xi_i\in W_i,\ i=1,$$

$2,\ \cdots,\ r\}$．容易证明 $\sum\limits_{i=1}^{r}W_i$是V的一个子空间，叫做W_1，W_2，…，

W_r 的和（空间）.

例 7　考虑$D^{(3)}$，$\varepsilon_1=(1,\ 0,\ 0)$，$\varepsilon_2=(0,\ 1,\ 0)$，$\varepsilon_3=(0,\ 0,\ 1)$，于是

$$D^{(3)}=L(\varepsilon_1,\varepsilon_2,\varepsilon_3)=L(\varepsilon_1,\ \varepsilon_2)+L(\varepsilon_3)=L(\varepsilon_1)+L(\varepsilon_2)+ +L(\varepsilon_3).$$

例 8　考虑：$M_2(F)$，$\varepsilon_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix}$，$\varepsilon_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$，

$\varepsilon_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix}$，$\varepsilon_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$. 于是

$$\begin{aligned}M_2(F)&=L(\varepsilon_{11},\ \varepsilon_{12},\ \varepsilon_{21},\ \varepsilon_{22})=L(\varepsilon_{11},\ \varepsilon_{12})+L(\varepsilon_{21},\ \varepsilon_{22})\\&=L(\varepsilon_{11})+L(\varepsilon_{12})+L(\varepsilon_{21})+L(\varepsilon_{22})\\&=L(\varepsilon_{11},\ \varepsilon_{12}+\varepsilon_{21},\ \varepsilon_{22})+L(\varepsilon_{12}-\varepsilon_{21})\end{aligned}$$

象例 6 那样的和子空间极为重要，为此我们引出以下概念.

定义 1　设W_1,W_2是线性空间V的两个子空间，如果 $W_1\cap W_2=\{0\}$,则称和W_1+W_2是直接和，简称为直和，记作$W=W_1\oplus W_2$. 这时也说W分解为W_1与W_2的直和.

前面例 6 中的子空间 S 与 T 的和是直和，即 $M_2(F)$ 分解为子空间 S 与 T 的直和：$M_2(F)=S\oplus T$. 例 7 中的子空间 $L(\varepsilon_1,\ \varepsilon_2)$ 与 $L(\varepsilon_3)$ 的和是直和，即 $D^{(3)}=L(\varepsilon_1,\ \varepsilon_2)\oplus L(\varepsilon_3)$。例 8 中的子空间 $L(\varepsilon_{11},\ \varepsilon_{12})$ 与子空间 $L(\varepsilon_{21},\ \varepsilon_{22})$ 的和是直和，即 $M_2(F)=L(\varepsilon_{11},\ \varepsilon_{12})\oplus L(\varepsilon_{21},\ \varepsilon_{22})$.

命题 1　和空间$W=W_1+W_2$是直和当且仅当和空间W中每一向量α的表示（分解）式

$$\alpha=\xi+\eta,\ \xi\in W_1,\ \eta\in W_2 \qquad (1)$$

是唯一的，即

如果 $\alpha=\xi+\eta=\xi'+\eta'$，$\xi$，$\xi'\in W_1$，$\eta$，$\eta'\in W_2$，那么必有 $\xi'=\xi$，$\eta'=\eta$.

证明　假设 $W=W_1+W_2$ 是直和，即 $W_1\cap W_2=\{0\}$，任取 $\alpha\in W$，如果 $\alpha=\xi+\eta=\xi'+\eta'$，$\xi$，$\xi'\in W_1$，$\eta$，$\eta'\in W_2$. 那么经

移项可有

$$\xi-\xi'=\eta'-\eta.$$

其中 $\xi-\xi'\in W_1$，$\eta'-\eta\in W_2$。从而 $\xi-\xi'\in W_2$，即得$\xi-\xi'\in W_1\cap W_2$，所以 $\xi-\xi'=0$。即 $\xi=\xi'$。自然也有 $\eta'=\eta$。因之，α 的表示式是唯一的。

反之，假设 $W=W_1+W_2$ 中每一向量α表成（1）的方式是唯一的，任取 $\beta\in W_1\cap W_2$，于是

$$0=\beta+(-\beta),\beta\in W_1,\quad -\beta\in W_2$$

但又有 $0=0+0$，由零向量的表示式是唯一的， 可得 $\beta=0$。即 $W_1\cap W_2=\{0\}$，所以 $W=W_1+W_2$ 是直和。

命题 2　**$W=W_1+W_2$ 是直和当且仅当零向量的表示（分解）式是唯一的，即**

$$0=\xi+\eta \text{ 时必有 } \xi=0,\ \eta=0.$$

证明　必要性是命题 1 的特款，现在只须证明充分性，为此指出由零向量表示式的唯一性可推出任一向量的表示式都是唯一的即可。假设和W中任一向量α有表示式

$$\alpha=\xi+\eta=\xi'+\eta',\ \xi,\ \xi'\in W_1,\ \eta,\ \eta'\in W_2.$$

则有

$$0=(\xi-\xi')+(\eta-\eta'),\ \xi-\xi'\in W_1,\ \eta-\eta'\in W_2.$$

于是因为零向量的表示式是唯一的，即得

$$\xi-\xi'=0,\ \eta-\eta'=0,\ \text{即}\ \xi'=\xi,\ \eta'=\eta.$$

所以，向量α的表示式是唯一的，由命题 1 和$W=W_1+W_2$是直和。

命题 1 反映出直和的意义在于，如果线性空间V能分解成两个子空间W_1与W_2的直和，那么V的一切向量及其运算完全由子空间W_1与W_2的向量及其运算所唯一确定，换句话说就是线性空间V的结构由两个子空间W_1与W_2的结构所决定，命题 2 可以做为直和简单判别法，用起来比较方便。

直和也容易推广为任意有限个子空间的情形。

定义 2　**设W_1，W_2，$\cdots$，W_r 是线性空间V的子空间，如果**

$W_i\cap(W_1+\cdots+W_{i-1}+W_{i+1}+\cdots+W_r)=\{0\}, 1\leqslant i\leqslant r$，则称和 $W=W_1+W_2+\cdots+W_r$ 是直接和，简称直和，记作$W=W_1\oplus W_2\oplus\cdots\oplus W_r$ 或 $W=W_1\oplus W_2\oplus\cdots\oplus W_r$。这时也说$W$分解为子空间$W_1$，$W_2$，…，$W_r$的直和。

与命题1、命题2的证明完全一样可以证明如下两个命题。

命题3　和$W=\sum_{i=1}^{r}W_i$ 是直和当且仅当W中每一向量α的表示（分解）式

$$\alpha=\xi_1+\xi_2+\cdots+\xi_r,\quad \xi_i\in W_i,\quad i=1,\ 2,\ \cdots,\ r$$

都是唯一的，即如果还有

$$\alpha=\eta_1+\eta_2+\cdots+\eta_r,\quad \eta_i\in W_i,\quad i=1,\ 2,\ \cdots,\ r$$

那么必有　$\xi_1=\eta_1$，$\xi_2=\eta_2$，…，$\xi_r=\eta_r$。

命题4　和 $W=\Sigma W_i$ 是直和当且仅当零向量的表示（分解）式是唯一的，即

$$0=\xi_1+\xi_2+\cdots+\xi_r\ \text{时},\ \xi_i\in W_i,\quad i=1,\ 2,\ \cdots,\ r$$

必有 $\xi_1=\xi_2=\cdots=\xi_r=0$。

这两个命题的具体证明做为练习留给读者，这两个命题的意义也是与命题1、2完全相当的。

例9　考虑 $F^{(3)}$，$\varepsilon_1=(1,\ 0,\ 0)$，$\varepsilon_2=(0,\ 1,\ 0)$，$\varepsilon_3=(0,\ 0,\ 1)$ 于是由命题4显然有

$$F^{(3)}=L(\varepsilon_1)\oplus L(\varepsilon_2)\oplus L(\varepsilon_3)$$

这表明三数组做成的三维向量空间可分解为三个一维子空间的直和，用几何语言来说就是三维几何空间可分解为三个坐标轴的直和，一般的，三维几何空间可分解为三条不共面的直线的直和。

例10　考虑 $F_n[x]$，于是由命题4易知

$$F_n[x]=L(1,x,x^2,\ \cdots,\ x^{n-1})=L(1)\oplus L(x)\oplus L(x^2)\oplus\cdots+L(x^{n-1})$$

这表明由一切次数低于n的多项式做成的线性空间可分解为n个子空间的直和，这n个子空间依次是由单项式 1，x，x^2，…，x^{n-1} 生成

的。

例11　考虑 $M_3(F)$。

令

$$\varepsilon_{ij}=\begin{pmatrix} & \vdots & \\ \cdots & 1 & \cdots \\ & \vdots & \end{pmatrix}\begin{matrix}\\ (i)\\ \\ \end{matrix}$$

(j)

(i)，未标出的位置，元素都是零，即

$$\varepsilon_{11}=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix},\quad \varepsilon_{12}=\begin{pmatrix}0&1&0\\0&0&0\\0&0&0\end{pmatrix},\quad \varepsilon_{13}=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix},$$

$$\varepsilon_{21}=\begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix},\quad \varepsilon_{22}=\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix},\quad \varepsilon_{23}=\begin{pmatrix}0&0&0\\0&0&1\\0&0&0\end{pmatrix},$$

$$\varepsilon_{31}=\begin{pmatrix}0&0&0\\0&0&0\\1&0&0\end{pmatrix},\quad \varepsilon_{32}=\begin{pmatrix}0&0&0\\0&0&0\\0&1&0\end{pmatrix},\quad \varepsilon_{33}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&1\end{pmatrix}.$$

于是由命题 4，则有

$$M_3(F)=L(\varepsilon_{11},\ \varepsilon_{12},\ \varepsilon_{13})\oplus L(\varepsilon_{21},\ \varepsilon_{22},\ \varepsilon_{23})\oplus L(\varepsilon_{31},\ \varepsilon_{32},\ \varepsilon_{33})$$

其中

$$L(\varepsilon_{11},\ \varepsilon_{12},\ \varepsilon_{13})=\left\{\begin{pmatrix}a_{11}&a_{12}&a_{13}\\0&0&0\\0&0&0\end{pmatrix}\middle|\ a_{11},\ a_{12},\ a_{13}\in F\right\}$$

$$L(\varepsilon_{21},\ \varepsilon_{22},\ \varepsilon_{23})=\left\{\begin{pmatrix}0&0&0\\a_{21}&a_{22}&a_{23}\\0&0&0\end{pmatrix}\middle|\ a_{21},\ a_{22},\ a_{23}\in F\right\}$$

$$L(\varepsilon_{31},\ \varepsilon_{32},\ \varepsilon_{33})=\left\{\begin{pmatrix}0&0&0\\0&0&0\\a_{31}&a_{32}&a_{33}\end{pmatrix}\middle|\ a_{31},\ a_{32},\ a_{33}\in F\right\}$$

当然还有进一步的分解式：

$$M_3(F)=L(\varepsilon_{11})\oplus L(\varepsilon_{12})\oplus L(\varepsilon_{13})\oplus\cdots\oplus L(\varepsilon_{31})\oplus L(\varepsilon_{32})\oplus L(\varepsilon_{33})。$$

练　习　三

1．举例说明两个子空间的并不必是子空间。

2．设 S_1，S_2 是V的子空间，证明：

1) $S_1+(S_1\cap S_2)=S_1$

2) $S_1\cap(S_1+S_2)=S_1$

3．设 W_1，W_2，W_3 是线性空间V的子空间，证明：

1) $W_1\cap[(W_1\cap W_2)+W_3]=(W_1\cap W_2)+(W_1\cap W_3)$

2) $W_1+[(W_1+W_2)\cap W_3]=(W_1+W_2)\cap(W_1+W_3)$

4．设 W_1，W_2，W_3是线性空间V的子空间，证明：

1) $(W_1\cap W_2)+W_3\subseteq(W_1+W_3)\cap(W_2+W_3)$

2) $(W_1\cap W_3)+(W_2\cap W_3)\subseteq(W_1+W_2)\cup W_3$

5．设W，W_1，W_2是线性空间V的子空间，其中$W_1\subseteq W_2$，如果

$$W\cap W_1=W\cap W_2$$
$$W+W_1=W+W_2$$

证明：$W_1=W_2$。

6．在线性空间V中，证明：所有包含子空间R和L的子空间的交等于$R+L$。即$R+L$是包含子空间R和L的最小子空间。

7．设W_1，W_2，…，W_n 是线性空间V的子空间，如果

$$W_i\cap W_j=\{0\}\qquad i\neq j,\ i,\ j=1,\ 2,\ \cdots,\ n$$

问等式

$$W_1+W_2+\cdots+W_n=W_1\oplus W_2\oplus\cdots\oplus W_n$$

是否成立？

§4 线性相关性

在以数组为向量做成的向量空间 $F^{(n)}$ 的理论中，线性相关性是个基本的重要概念，它在解决线性方程组解的结构问题中所起的重要作用是大家所熟悉的，对一般的线性空间理论来说，线性相关性同样是一个十分基本的重要概念，我们就是依据向量的线性运算，通过线性相关性的概念分析向量之间的普遍联系，从而解决线性空间的基本课题。

虽然一般线性空间中线性相关性的定义及其基本性质与以数组为向量的向量空间中相应的定义及性质几乎完全是一致的。但是为了达到复习的目的和相互做些对照以加强记忆和理解，以下我们还是完整的把它们重述出来。

设V是F上任一线性空间。

定义 1 设α_1，α_2，…，α_r，$\beta\in V$。如果有k_1，k_2，…，$k_r\in F$，使

$$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r \tag{1}$$

成立，则称β是α_1，α_2，…，α_r的线性组合，或者说β能用α_1，α_2，…，α_r线性表出，k_1，k_2，…，k_r叫做表出系数。

定义 2 设

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r \tag{2}$$

$$\beta_1,\beta_2,\ \cdots,\ \beta_s \tag{3}$$

是V中两组向量，如果组(2)中每一向量都能用组(3)线性表出，则称向量组(2)能用向量组(3)线性表出，如果向量组(2)能用向量组(3)线性表出，同时向量组(3)也能用向量组(2)线性表出，则称向量组(2)与向量组(3)等价。

定义 3 设α_1，α_2，…，α_r是V中一组向量，如果F中有不全为零的r个数k_1，k_2，…，k_r使

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r=0 \tag{4}$$

成立，则称 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是线性相关的；如果只有 $k_1=k_2=\cdots=k_r=0$ 时(4)式才成立，则称 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是线性无关的。

定义4　设S是V的一个子集，S中的一组向量 $\alpha_1, \alpha_2, \cdots, \alpha_r$，如果

1) $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是线性无关的。

2) 对S中任意向量β，向量组 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta$ 都是线性相关的，则称 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是S的一个极大线性无关组。

以上对一般线性空间叙述了与线性相关性有关的几个基本概念，它们在线性空间的讨论中随时都要用到，显然，在第六章中对于n维向量空间定义的相应概念是这里一般概念的特殊情形，下面看几个简单的例子。

例1　$L(\alpha_1, \alpha_2, \cdots, \alpha_r)$ 恰好是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的一切可能的线性组合组成的。

例2　$L(\alpha_1, \alpha_2, \cdots, \alpha_r)=L(\beta_1, \beta_2, \cdots, \beta_t)$ 当且仅当 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 与 $\beta_1, \beta_2, \cdots, \beta_t$ 等价。

例3　考虑 $F[x]$。$1, x, x^2, \cdots, x^{n-1}\in F_n[x]$，于是

1) $1, x, x^2, \cdots, x^{n-1}$ 是线性无关的，

2) 对任一 $f(x)\in F_n[x]$，都有

$$1, x, x^2, \cdots, x^{n-1}, f(x)$$

是线性相关的。

事实上欲使 $a_0+a_1x+a_2x^2+\cdots+a_{n-1}x^{n-1}=0$ 成立，按多项式等于零的定义，只有各个系数 $a_0, a_1, a_2, \cdots, a_{n-1}$ 全为零才行，所以 $1, x, x^2, \cdots, x^{n-1}$ 是线性空间 $F[x]$ 中一组线性无关的向量。

再有 $f(x)\in F_n[x]$，可令 $f(x)=c_0+c_1x+c_2x^2+\cdots+c_{n-1}x^{n-1}$，于是

$$c_0+c_1x+c_2x^2+\cdots+c_{n-1}x^{n-1}+(-1)f(x)=0$$

上式中诸系数 $c_0, c_1, c_2, \cdots, c_{n-1}, -1$ 不全为0，按定义3，$1, x, x^2, \cdots, x^{n-1}, f(x)$ 是线性相关的。

由于 1, x, x^2, …, x^{n-1} 是子集$F_n[x]$ 的一组向量，1)与2)说明$1, x, x^2$, …, x^{n-1}是$F_n[x]$的一个极大线性无关组。

例4 V是任一线性空间，$\alpha \in V$。于是α线性相关当且仅当$\alpha = 0$。

命题1 一组向量 α_1, α_2, …, α_s ($s \geqslant 2$) 是线性相关的当且仅当当中只少有一个向量能用其余向量线性表出。

命题2 若α_1, α_2, …, α_r 线性无关，α_1, α_2, …, α_r, β线性相关，那么β能用 α_1, α_2, …, α_r 线性表出，并且表法唯一。

以上两个命题的证明与第六章相应命题的证明完全一样，建议读者做为练习，把这两个证明再重写一遍。

命题3 如果α_1, α_2, …, α_r 能用 β_1, β_2, …, β_s 线性表出，而且 $r > s$，那么 α_1, α_2, …, α_r 线性相关。

证明 考察线性表出式

$$0 = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_r\alpha_r \tag{5}$$

由已知条件，α_1, α_2, …, α_r 能用 β_1, β_2, …, β_s 线性表出，可令

$$\begin{cases} \alpha_1 = a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1s}\beta_s \\ \alpha_2 = a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2s}\beta_s \\ \cdots \quad \cdots \quad \cdots \\ \alpha_r = a_{r1}\beta_1 + a_{r2}\beta_2 + \cdots + a_{rs}\beta_s \end{cases} \tag{6}$$

把（6）代入（5），即有

$$\begin{aligned} 0 &= k_1(a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1s}\beta_s) + k_2(a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2s}\beta_s) + \\ &\quad + \cdots + k_r(a_{r1}\beta_1 + a_{r2}\beta_2 + \cdots + a_{rs}\beta_s) \\ &= (a_{11}k_1 + a_{21}k_2 + \cdots + a_{r1}k_r)\beta_1 + (a_{12}k_1 + a_{22}k_2 + \cdots + a_{r2}k_r)\beta_2 + \\ &\quad + \cdots + (a_{1s}k_1 + a_{2s}k_2 + \cdots + a_{rs}k_r)\beta_s. \end{aligned}$$

如果有k_1, k_2, …, k_r 使

$$\begin{cases} a_{11}k_1 + a_{21}k_2 + \cdots + a_{r1}k_r = 0 \\ a_{12}k_1 + a_{22}k_2 + \cdots + a_{r2}k_r = 0 \\ \cdots \quad \cdots \quad \cdots \\ a_{1s}k_1 + a_{2s}k_2 + \cdots + a_{rs}k_r = 0 \end{cases} \tag{7}$$

成立，那么这r个数也使（5）成立，这里（7）是一个齐次线性方程组，当 $r>s$ 时，必有非零解．于是任取（7）的一个非零解 $k_1, k_2, \cdots, k_r$．将其代入（5）式，就证明 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是线性相关的．

推论 1 若 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 能用 $\beta_1, \beta_2, \cdots, \beta_s$ 线性表出，而且 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关，那么 $r \leqslant s$．

推论 2 若 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 与 $\beta_1, \beta_2, \cdots, \beta_s$ 都线性无关而且是等价的，那么 $r=s$．

推论 3 如果子集S有极大线性无关组，那么它的各个极大线性无关组中所含向量的个数都相等．

定义 5 如果 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是S的极大线性无关组，则称r是S的秩．特别地当S中只含零向量时，称S的秩等于 0．

显然，等价的向量组秩一定相等，但反之则未必．

命题 4 V中任意有限个不全为零的向量

$$\alpha_1, \alpha_2, \cdots, \alpha_s$$

必有极大线性无关组，从而必有秩．

证明 不妨认为 $\alpha_1 \neq 0$．于是观察

$$\alpha_1, \alpha_2, \cdots, \alpha_s$$

从左边开始，依次去掉那些能用它前面各向量线性表出的一切向量、记剩下的向量为

$$\alpha_{i_1}(=\alpha_1), \alpha_{i_2}, \cdots, \alpha_{i_r}, \quad i_1(=1)<i_2<\cdots<i_r.$$

容易指出，上述剩下的这r个向量就是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的一个极大线性无关组．

首先，$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 是线性无关的．不然，当中必有能用它前面的向量线性表出的向量，这与该向量被留下相矛盾，所以 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 必是线性无关的．

其次，对 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 中任一向量α_l

$$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}, \alpha_l \tag{8}$$

必线性相关．这可分两种情形来说明，若α_l是 $\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$ 中

的一个．（8）当然是线性相关的．若α_l不是α_{i_1}，α_{i_2}，…，α_{i_r}中的一个，则必有

$$\alpha_{i_1}, \cdots, \alpha_{i_k}, \alpha_l, \alpha_{i_{k+1}}, \cdots, \alpha_s, \quad i_k<l<i_{k+1}.$$

因为α_l不是α_{i_1}，α_{i_2}，…，α_{i_r}中的任何一个，这就是说α_l是被去掉的一个向量．从而α_l能用它前边的向量$\alpha_{i_1},\cdots$，α_{i_k}线性表出．所以α_{i_1}，…，α_{i_k}，α_l线性相关，因此向量组α_{i_1}，…，α_{i_k}，α_l，是线性相关的．这就证明α_{i_1}，α_{i_2}，…，α_{i_r}是α_1，α_2，…，α_s的一个极大线性无关组．

推论1 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ **有极大线性无关组，特别地，**α_1，α_2，…，α_s **的极大线性无关组都是** $\mathbf{L}(\alpha_1, \alpha_2, \cdots, \alpha_s)$ **的极大线性无关组．**

推论2 $L(\alpha_1, \alpha_2, \cdots, \alpha_s)$ **中任何一组线性无关向量**β_1，β_2，…，β_t **都可以扩充成为** $\mathbf{L}(\alpha_1,\alpha_2, \cdots, \alpha_s)$ **的极大线性无关组．**

事实上 考虑

$$\beta_1, \beta_2, \cdots, \beta_t, \alpha_1, \alpha_2, \cdots, \alpha_s.$$

按命题4的证法，从左往右去掉能用它前面诸向量线性表出的一切向量．由于β_1，β_2，…，β_t是线性无关的，不可能有任何一个β_i被去掉．这样剩下的向量为

$$\beta_1, \beta_2, \cdots, \beta_t, \alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{j_l}$$

显然，这就是从t个线性无关的向量β_1，β_2，…，β_t的基础上扩充成的一个极大线性无关组．

例5 线性空间$F[x]$的子集$S=\{1, x, x^2, \cdots, x^m, \cdots\}$没有极大线性无关组．

事实上这是比较明显的．如果S有极大线性无关组，令其秩为r．于是S中多于r个向量的向量组必线性相关．但是我们知道，S中任意有限个向量总是线性无关的．所以S没有极大线性无关组．

练 习 四

1．设 α_1，α_2，…，α_r是线性空间V中线性无关的向量组，讨论 $k_1\alpha_1$，$k_2\alpha_2$，…，$k_r\alpha_r$ 的线性相关性。

2．在实函数线性空间中，证明：$1, \cos^2 t, \cos 2t$是线性相关的。

3．在 $F_n[x]$ 中，证明：

$$1,\ 1+x,\ \cdots,\ 1+x+\cdots+x^{n-1}$$

是线性无关的。

4．如果 $f_1(x)$，$f_2(x)$，$f_3(x)$ 是线性空间 $F[x]$ 中三个互质多项式，其中任意两个都不互质，那么它们必线性无关。

5．设 α_1，α_2，…，α_r 是一组线性无关的向量，任取 k_1，k_2，…，k_{r-1} 这 $r-1$ 个数，令

$$\beta_1=\alpha_1+k_1\alpha_r$$
$$\beta_2=\alpha_2+k_2\alpha_r$$
$$\cdots\cdots$$
$$\beta_{r-1}=\alpha_{r-1}+k_{r-1}\alpha_r$$
$$\beta_r=\alpha_r$$

那么 β_1，β_2，…，β_r 是线性无关的。

6．设向量β可被 α_1，α_2，…，α_r 线性表出，但β 不能被 α_1，α_2，…，α_{r-1} 线性表出，证明：向量组 α_1，α_2，…，α_r 与向量组 α_1，α_2，…，α_{r-1}，β 等价。

7．设向量组 α_1，α_2，…，α_r 线性无关，而α_1，α_2，…α_r，β，γ 线性相关，证明：或β与γ中至少有一个可以由α_1，α_2，…，α_r 线性表出，或向量组 $\alpha_1,\alpha_2,\cdots,\alpha_r$，$\beta$ 与向量组 α_1，α_2，…，α_r，γ等价。

§5 有限维线性空间

与 n维向量空间的情形一样，极大线性无关组对一般线性空间

的结构的重要性也是明显的。比如线性空间V有极大线性无关组，而且秩等于n。那么任意取定V的一个极大线性无关组ε_1，ε_2，…，ε_n，对$\forall\xi\in V$，都有

$$\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n, \quad x_i \in F.$$

其中表出系数x_1，x_2，…，x_n是由向量ξ所唯一决定的，进而若$\eta\in V$，而

$$\eta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n$$

则有

$$\xi+\eta = (x_1+y_1)\varepsilon_1 + (x_2+y_2)\varepsilon_2 + \cdots + (x_n+y_n)\varepsilon_n,$$

$$k\xi = (kx_1)\varepsilon_1 + (kx_2)\varepsilon_2 + \cdots + (kx_n)\varepsilon_n.$$

这就表明，线性空间V的每一向量都由V的极大线性无关组——有限个向量ε_1，ε_2，…，ε_n以唯一的一组系数线性表示出来；并且V中向量的基本运算：加法与倍数也完全通过它们用ε_1，ε_2，…，ε_n的表出系数来进行相应的运算。这等于说，对于有极大线性无关组的线性空间V来说，V中的向量（一定是无限多个！）及其运算可以通过有限个向量来实现。当然这样的线性空间的结构是比较简单易于掌握的。但是与以数组为向量的向量空间不同，对一般的线性空间来说，并不都有极大线性无关组，如上节最后的例表明，线性空间$F[x]$就没有极大线性无关组。本节专门讨论有极大线性无关组的线性空间的基本性质。

定义 **设V是F上的线性空间，如果V有极大线性无关组，则称V是F上的有限维空间。V的每一个极大线性无关组都叫做V的一个基底。极大线性无关组所含向量个数n叫做V的维数，记作$\dim V=n$。这时就说V是n维空间。我们约定，零空间也叫有限维空间，其维数为零。若V不是有限维空间，则称V是无限维空间，有时用$V_n(F)$表示V是F上n维空间。**

如果$\dim V=n$，ε_1，ε_2，…，ε_n是V的任一基底，这时就记作$V=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$。

例1　在$F^{(3)}$中，$\varepsilon_1=(1, 0, 0)$，$\varepsilon_2=(0,1,0)$，$\varepsilon_3=(0,0,1)$。

则有 $F^{(3)}=[\varepsilon_1, \varepsilon_2, \varepsilon_3]$。

一般地，$F^{(n)}=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$，其中 $\varepsilon_i=(0, \cdots, 0, \overset{(i)}{1}, 0, \cdots, 0)$，$i=1, 2, \cdots, n$。

例 2　$F_n[x]=[1, x, x^2, \cdots, x^{n-1}]$，即 $\dim F_n[x]=n$。$1, x, x^2, \cdots, x^{n-1}$ 是 $F_n[x]$ 的一个基底。

例 3　$M_2(F)=[\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{21}, \varepsilon_{22}]$，其中

$$\varepsilon_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix}, \varepsilon_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix}, \varepsilon_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix}, \varepsilon_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix},$$

即 $\dim M_2(F)=4$，$\varepsilon_{11}, \varepsilon_{12}, \varepsilon_{21}, \varepsilon_{22}$ 是$M_2(F)$ 的一个基底。

例 4　$W=\{A\in M_2(F) \mid A'=A\}$。取

$$\delta_1=\begin{pmatrix}1&0\\0&0\end{pmatrix}, \delta_2=\begin{pmatrix}0&0\\0&1\end{pmatrix}, \delta_3=\begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

则有

$$\begin{pmatrix}a&c\\c&b\end{pmatrix}=a\delta_1+b\delta_2+c\delta_3.$$

已知 $W=L(\delta_1, \delta_2, \delta_3)$。又知 $\delta_1, \delta_2, \delta_3$ 是线性无关的，即 $\delta_1, \delta_2, \delta_3$ 是W的一个基底，所以，$W=[\delta_1, \delta_2, \delta_3]$，$\dim W=3$。

例 5　$F[x]$ 是无限维空间。

命题 1　设 $\dim V=n$。那么V中任意 n 个线性无关的向量必是基底，任意n个生成向量必是基底。

证明　由V的维数为n可知V中任一极大线性无关组都含n个向量。这样，如果$\alpha_1, \alpha_2, \cdots, \alpha_n$ 是V中线性无关的n个向量，那么，$\alpha_1, \alpha_2, \cdots, \alpha_n$ 必已是极大线性无关组。所以$\alpha_1, \alpha_2, \cdots, \alpha_n$ 就是V的一个基底。再如 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是n个生成向量，即$V=L(\alpha_1, \alpha_2, \cdots, \alpha_n)$。这时 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 必为线性无关的。不然，从$\alpha_1, \alpha_2, \cdots, \alpha_n$ 去掉若干个（至少得去掉一个！）向量之后，可得出V的一个极大线性无关组 $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{im}$，而且 $m<n$。这是不可能的。所以n个生成向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 必线性无关，从而是一个极大线性无关组，即 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是V的一个基底。

命题 2 设$V=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$. $\alpha_1, \alpha_2, \cdots \alpha_r$ $(r<n)$ 为一组线性无关向量，那么 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 可以扩充成V的一个基底.

此命题 2 就是上节命题 4 推论 2 的特例.

命题 3 V是有限维的当且仅当V是有限生成的.

证明 必要性是自明的. 充分性也是容易的. 因为$V=L(\alpha_1, \alpha_2, \cdots, \alpha_s)$，那么，$\alpha_1, \alpha_2, \cdots, \alpha_s$ 这有限个向量必有极大线性无关组 $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{ir}$ 它也就是空间V的一个基底，所以V是有限维的，而且 $\dim V=r=\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ 的秩.

命题 4 设V是有限维的，W是V的子空间，那么

1° W是有限维的，且

$\dim W\leqslant\dim V$，只在 $W=V$ 时才有 $\dim W=\dim V$.

2° 存在子空间K，使 $V=W\oplus K$

证明 1° 设 $\dim V=n$. 于是W 中任一组向量，如果所含向量个数多于 n 个，则必线性相关. 所以W必有极大线性无关组. 故W是有限维的. 而 $\dim W\leqslant\dim V$ 是显然的. 当 $\dim W=\dim V$ 时，W的基底也是V的基底，所以$W=V$.

2° 把 $W=[\varepsilon_1, \cdots, \varepsilon_r]$ 的基底扩充成V的一个基底 $\varepsilon_1, \cdots, \varepsilon_r, \delta_1, \cdots, \delta_{n-r}$. 令

$$K=[\delta_1, \delta_2, \cdots, \delta_{n-r}]$$

则有

$$V=W+K$$

因为 $\varepsilon_1, \cdots, \varepsilon_r, \delta_1, \delta_2, \cdots, \delta_{n-r}$ 线性无关，所以 $W+K$ 是直和，即 $V=W\oplus K$.

命题 5 $\dim(W_1+W_2)=\dim W_1+\dim W_2-\dim(W_1\cap W_2)$

证明 在子空间$W_1\cap W_2$ 中取定一个基底：$\alpha_1, \alpha_2, \cdots \alpha_r$. 因为 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 是 W_1 的一组线性无关向量. 所以能把它扩充成为W_1的一个基底：$\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_1, \beta_2, \cdots, \beta_s$; 同样地，把 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 扩充成为W_2的一个基底：$\alpha_1, \alpha_2, \cdots, \alpha_r, \gamma_1, \gamma_2, \cdots, \gamma_t$

考察向量组

$$\alpha_1,\ \alpha_2\ \cdots,\ \alpha_r,\ \beta_1,\ \beta_2,\ \cdots,\ \beta_s,\ \gamma_1,\ \gamma_2,\ \cdots,\ \gamma_t.$$

显然这是 W_1+W_2 中的一组向量。我们指出它是 W_1+W_2 的一个基底。

首先，这 $r+s+t$ 个向量是 W_1+W_2 的一组生成元素，这是十分明显的。

其次，令

$$a_1\alpha_1+a_2\alpha_2+\cdots+a_r\alpha_r+b_1\beta_1+b_2\beta_2+\cdots+b_s\beta_s+c_1\gamma_1+c_2\gamma_2+ \\ +\cdots+c_t\gamma_t=0$$

如果 $c_1\gamma_1+c_2\gamma_2+\cdots+c_t\gamma_t\neq 0$，由

$$c_1\gamma_1+c_2\gamma_2+\cdots+c_t\gamma_t=-(a_1\alpha_1+a_2\alpha_2+\cdots+a_r\alpha_r+b_1\beta_1+ \\ +b_2\beta_2+\cdots+b_s\beta_s)$$

可知 $c_1\gamma_1+c_2\gamma_2+\cdots+c_t\gamma_t\in W_1\cap W_2$。从而应有 $c_1\gamma_1+c_2\gamma_2+\cdots+c_t\gamma_t$ 能用 α_1，α_2，$\cdots$，α_r 线性表出：

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r=c_1\gamma_1+c_2\gamma_2+\cdots+c_t\gamma_t,$$

而且 k_1，k_2，$\cdots$，k_r 不全为0。

于是

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r+(-c_1)\gamma_1+(-c_2)\gamma_2+ \\ +\cdots+(-c_t)\gamma_t=0.$$

这与 α_1，α_2，$\cdots$，α_r，γ_1，γ_2，$\cdots$，γ_t 线性无关相矛盾。这就说明，$c_1\gamma_1+c_2\gamma_2+\cdots+c_t\gamma_t=0$。

因为 γ_1，γ_2，$\cdots$，γ_t 线性无关，所以 $c_1=0, c_2=0, \cdots, c_t=0$。进而就有

$$a_1\alpha_1+a_2\alpha_2+\cdots+a_r\alpha_r+b_1\beta_1+b_2\beta_2+\cdots+b_s\beta_s=0.$$

这样又得出 $a_1=0$，$a_2=0$，$\cdots$，$a_r=0$，$b_1=0$，$b_2=0$，$\cdots$，$b_s=0$。总之，即得 α_1，α_2，$\cdots$，α_r，β_1，β_2，$\cdots$，β_s，γ_1，γ_2，$\cdots$，γ_t 是线性无关的。于是

$$\dim(W_1+W_2)=r+s+t,\qquad \dim W_1=r+s,$$
$$\dim W_2=r+t,\qquad \dim(W_1\cap W_2)=r.$$

所以 $\dim(W_1+W_2)=\dim W_1+\dim W_2-\dim(W_1\cap W_2)$

推论 W_1+W_2 是直和当且仅当 $\dim(W_1+W_2)=\dim W_1+\dim W_2$。

练 习 五

1．下列向量组是否是$F_4[x]$ 的基底

1）$\alpha_1=x^3+1$，$\alpha_2=x+1$，$\alpha_3=x^2+x$，$\alpha_4=x^3+x^2+2x+2$

2）$\beta_1=1-x^2$，$\beta_2=x-1$，$\beta_3=x^2+2x-2$，$\beta_4=x^3$

2．把向量组 $\alpha_1=(2, 1, -1, 3)$，$\alpha_2=(-1, 0, 1, 2)$扩充为$D^{(4)}$ 的基底。

3．设 S 是数域 F 上一切 n 阶对称阵所构成的线性空间，求 S 的维数。

4．证明：复数域C 做为实数域 D 上的线性空间，其维数是 2．如果C 看作自身上的线性空间，其维数是多少？

5．证明：如果线性空间中每个向量都可唯一地表成 α_1，α_2，…，α_n 的线性组合，证明该线性空间的维数是 n。

6．若 α_1，α_2，…，α_n 是线性空间V的生成组，且对 V 中某一向量β表法唯一，则 α_1，α_2 …，α_n 是V的基底。

7．设W是$D^{(n)}$ 的一个非零子空间，如果W中每个向量 $(a_1, a_2, \cdots, a_n)$ 要末 $a_1=a_2=\cdots=a_n=0$，要末每个 a_i 都不为0，证明：$\dim W=1$

§6 坐　　标

设V是n维线性空间，取定一个基底：

$$\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n.$$

于是对任意的 α，$\beta\in V$，都有唯一的表示式

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n,\qquad \beta=b_1\varepsilon_1+b_2\varepsilon_2+\cdots+b_n\varepsilon_n$$

而且

$$\alpha+\beta=(a_1+b_1)\varepsilon_1+(a_2+b_2)\varepsilon_2+\cdots+(a_n+b_n)\varepsilon_n,$$
$$k\alpha=(ka_1)\varepsilon_1+(ka_2)\varepsilon_2+\cdots+(ka_n)\varepsilon_n.$$

对此我们给出以下重要的

定义 V中向量α用基底 ε_1, ε_2, $\cdots$, ε_n 线性表出的 n 个表出系数叫做α在基底 ε_1, ε_2, $\cdots$, ε_n 上的坐标。

例如，$F_3[x]=[1, x, x^2]$。于是向量 $f(x)=1+2x+3x^2$ 在基底 1, x, x^2 上的坐标就是它的三个系数1, 2, 3。另外，$F_3[x]=[x, x^2, 1]$，这样，$f(x)=1+2x+3x^2$ 在这个基底上的坐标应当是2，3，1。这表明，在涉及坐标的讨论中，基底向量的次序是有关系的，一般来说，同样一组向量做成的基底，当向量次序不相同时，应当认为它们是不同的基底。如

1, x, x^2; x, x^2, 1; x^2, 1, x 等等

就是 $F_3[x]$ 的不同基底。

为了书写方便，我们约定以下两种记法。

1）如果α的坐标为 a_1, a_2, $\cdots$, a_n，那么令

$$\bar{\alpha}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},$$ 叫做向量α的坐标列（矩阵）；

$\bar{\alpha}'=(a_1, a_2, \cdots, a_n)$ 叫做α的坐标行（矩阵）。

2）把线性表出式子 $x_1\xi_1+x_2\xi_2+\cdots+x_r\xi_r=\beta$ 记作

$$x_1\xi_1+x_2\xi_2+\cdots+x_r\xi_r=(\xi_1\xi_2\cdots\xi_r)\begin{pmatrix}x_1\\x_2\\\vdots\\x_r\end{pmatrix}$$

或者 $$x_1\xi_1+x_2\xi_2+\cdots+x_r\xi_r=(x_1x_2\cdots x_r)\begin{pmatrix}\xi_1\\\xi_2\\\vdots\\\xi_r\end{pmatrix}.$$

其中1）的记法就是把α的坐标做成 $n\times1$ 矩阵或 $1\times n$ 矩阵，而

2）是一种形式记法，但是这种记法，尽管当中的 $(\xi_1\xi_2\cdots\xi_r)$ 的元素 $\xi_1, \xi_2, \cdots, \xi_r$ 未必是数，理解成“矩阵”的乘法是完全可以的。

一般的，如果有一组线性表出的式子

$$
\begin{aligned}
\alpha_1 &= a_{11}\xi_1 + a_{21}\xi_2 + \cdots + a_{r1}\xi_r \\
\alpha_2 &= a_{12}\xi_1 + a_{22}\xi_2 + \cdots + a_{r2}\xi_r \\
&\cdots \quad \cdots \quad \cdots \\
\alpha_s &= a_{1s}\xi_1 + a_{2s}\xi_2 + \cdots + a_{rs}\xi_r
\end{aligned} \tag{1}
$$

把（1）中每一个向量 α_i 的表出系数依次做列构成一个 $r\times s$ 矩阵

$$
A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \cdots & \cdots & & \cdots \\ a_{r1} & a_{r2} & \cdots & a_{rs} \end{pmatrix}
$$

参照2）的记法，可把（1）中的 s 个线性表出式子记作

$$(\alpha_1\alpha_2\cdots\alpha_s) = (\xi_1\xi_2\cdots\xi_r)A$$

当 $\xi_1, \xi_2, \cdots, \xi_r$ 是基底时，把矩阵 A 叫做 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的坐标列矩阵。

对于这种形式记法，容易验证下面的运算规则成立。

$$((\xi_1\ \xi_2\ \cdots\xi_r)A)B = (\xi_1\xi_2\cdots\xi_r)(AB)$$

$$(\xi_1\ \xi_2\cdots\xi_r)(A+B) = (\xi_1\ \xi_2\ \cdots\xi_r)A + (\xi_1\ \xi_2\ \cdots\xi_r)B$$

$$(\xi_1\ \xi_2\ \cdots\xi_r)A + (\eta_1\ \eta_2\ \cdots\eta_r)A = (\xi_1+\eta_1\ \xi_2+\eta_2\cdots\xi_r+\eta_r)A$$

$$k(\xi_1\ \xi_2\ \cdots\xi_r)A = (k\xi_1\ k\xi_2\cdots k\xi_r)A = (\xi_1\ \xi_2\ \cdots\xi_r)kA.$$

设 V 是 F 上 n 维线性空间。$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 为任意取定的一个基底。$F^{(n)}$ 是 n 数组空间，这里把 $F^{(n)}$ 中的数组一律写成 $n\times 1$ 矩阵的形式。这样，V 与 $F^{(n)}$ 之间建立起一个一一对应关系，即对任意的

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n \quad 或 \quad \alpha = (\varepsilon_1\varepsilon_2\cdots\varepsilon_n)\overline{\alpha}$$

$$\varphi:\ \alpha \longmapsto \overline{\alpha}.$$

进而对 V 中任意一组向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 自然有数组空间 $F^{(n)}$ 中的

s 个 “列向量” $\bar{\alpha}_1$, $\bar{\alpha}_2$, …, $\bar{\alpha}_s$，再依次以 $\bar{\alpha}_1$, $\bar{\alpha}_2$, …, $\bar{\alpha}_s$ 做列便构成一个 $n\times s$ 矩阵 $A=(\bar{\alpha}_1\ \bar{\alpha}_2\ \cdots\bar{\alpha}_s)$——$\alpha_1$, α_2, …, α_s 的坐标列矩阵．它们之间的具体关系就是

$$(\alpha_1\ \alpha_2\ \cdots\alpha_s)=(\varepsilon_1\varepsilon_2\cdots\varepsilon_n)A. \qquad (2)$$

下面我们指出，向量组 α_1, α_2, …, α_s 的线性相关性与其坐标列 $\bar{\alpha}_1$, $\bar{\alpha}_2$, …, $\bar{\alpha}_s$ 的线性相关性是完全一致的，从而可以通过数组空间 $F^{(n)}$ 的线性相关性理论来具体解决一般 n 维空间 V 中的线性相关性问题．同时也就可以利用矩阵的秩数理论来解决有限维空间的线性相关性问题．

考察线性表出式

$$0=k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s \qquad (3)$$

把向量 α_1, α_2, …, α_s 被基底 ε_1, ε_2, …, ε_n 的表示式（2）代入关系式（3）则有

$$\begin{aligned}0=&k_1(a_{11}\varepsilon_1+a_{21}\varepsilon_2+\cdots+a_{n1}\varepsilon_n)\\&+k_2(a_{12}\varepsilon_1+a_{22}\varepsilon_2+\cdots+a_{n2}\varepsilon_n)\\&+\cdots+\\&+k_s(a_{1s}\varepsilon_1+a_{2s}\varepsilon_2+\cdots+a_{ns}\varepsilon_n)\end{aligned}$$

把 k_i 乘到每个括号里边去，按 $\varepsilon_1,\varepsilon_2$, …, ε_n 合并各项，即得

$$\begin{aligned}0=&(a_{11}k_1+a_{12}k_2+\cdots+a_{1s}k_s)\varepsilon_1\\&+(a_{21}k_1+a_{22}k_2+\cdots+a_{2s}k_s)\varepsilon_2\\&+\cdots+\\&+(a_{n1}k_1+a_{n2}k_2+\cdots+a_{ns}k_s)\varepsilon_n\end{aligned} \qquad (4)$$

因为 ε_1, ε_2, …, ε_n 是线性无关的，所以可得与（4）等效的式子：

$$\begin{cases}a_{11}k_1+a_{12}k_2+\cdots+a_{1s}k_s=0\\a_{21}k_1+a_{22}k_2+\cdots+a_{2s}k_s=0\\\cdots\quad\cdots\quad\cdots\\a_{n1}k_1+a_{n2}k_2+\cdots+a_{ns}k_s=0\end{cases} \qquad (5)$$

(5)式成立说明 $k_1, k_2, \cdots, k_s$ 是齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1s}x_s = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2s}x_s = 0 \\ \cdots \quad \cdots \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{ns}x_s = 0 \end{cases} \tag{6}$$

的解，这里（6）的系数阵恰好就是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的坐标列矩阵A。这样我们证明了

定理 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的坐标列矩阵为A。$k_1, k_2, \cdots, k_s$是F中的一组数，那么

$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0$ 成立当且仅当 $k_1, k_2, \cdots, k_s$ 是齐次方程组(6)的解当且仅当 $k_1, k_2, \cdots, k_s$ 是向量方程

$$k_1\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + k_2\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \cdots + k_s\begin{pmatrix} a_{1s} \\ a_{2s} \\ \vdots \\ a_{ns} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

的解。

如果利用上边引进的形式记法，不仅可以简化演算过程，而且能够集中体现出一般向量 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 与其坐标列向量$\overline{\alpha_1}, \overline{\alpha_2}, \cdots, \overline{\alpha_s}$线性关系的一致性。如

$$(\alpha_1\ \alpha_2\ \cdots\alpha_s)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_s \end{pmatrix} = 0 \text{ 当且仅当} A\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

当且仅当 $(\overline{\alpha_1}\ \overline{\alpha_2}\cdots\ \overline{\alpha_s})\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_s \end{pmatrix} = 0$。

推论 1 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关当且仅当 $\mathrm{rank}A < s$ 当且仅当$\overline{\alpha_1}, \overline{\alpha_2}, \cdots, \overline{\alpha_s}$线性相关，

推论 2 $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{ir}$ 是 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的极大线性无关

组当且仅当$\bar{\alpha}_{i1}$，$\bar{\alpha}_{i2}$，…，$\bar{\alpha}_{ir}$是$\bar{\alpha}_1$，$\bar{\alpha}_2$，…，$\bar{\alpha}_s$的极大线性无关组当且仅当$\bar{\alpha}_{i1}$，$\bar{\alpha}_{i2}$，…，$\bar{\alpha}_{ir}$上有A的基础子式。

总结以上讨论表明，对任一n维线性空间V，任意取定一个基底ε_1，ε_2，…，ε_n之后，便有

1）V中向量在基底ε_1，ε_2，…，ε_n上的坐标列恰好就是n数组空间$F^{(n)}$的向量。

2）V与$F^{(n)}$之间的映射

$$\varphi:\ \alpha \longmapsto \bar{\alpha}$$

是一个一一对应，其中$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n$，$\bar{\alpha} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$。

3）$\overline{\alpha+\beta} = \bar{\alpha} + \bar{\beta}$，

$\overline{k\alpha} = k\bar{\alpha}$。

4）V中一组向量α_1，α_2，…，α_s唯一确定$F^{(n)}$中s个n数组$\bar{\alpha}_1$，$\bar{\alpha}_2$，…，$\bar{\alpha}_s$，从而唯一确定一个$n\times s$矩阵A——坐标列矩阵：

$$\alpha_1, \alpha_2, \ \cdots, \ \alpha_s \longrightarrow \bar{\alpha}_1, \ \bar{\alpha}_2, \ \cdots, \bar{\alpha}_s \longrightarrow A = (\bar{\alpha}_1\ \bar{\alpha}_2 \cdots \bar{\alpha}_s)。$$

于是

α_1，α_2，…，α_s线性相关当且仅当$\bar{\alpha}_1$，$\bar{\alpha}_2$，…，$\bar{\alpha}_s$线性相关当且仅当$\operatorname{rank}A<s$。

特别地

α_{i1}，α_{i2}，…，α_{ir}是α_1，α_2，…，α_s的极大线性无关组当且仅当$\bar{\alpha}_{i1}$，$\bar{\alpha}_{i2}$，…，$\bar{\alpha}_{ir}$是$\bar{\alpha}_1$，$\bar{\alpha}_2$，…，$\bar{\alpha}_s$的极大线性无关组当且仅当α_{i1}，α_{i2}，…，α_{ir}的坐标列上有A的基础子式.

这样，一般n维线性空间V以及其中向量之间的线性关系就完全的具体的转化为n数组空间$F^{(n)}$及其中相应向量之间的线性关系，因而n数组空间的理论对有限维线性空间来说具有代表性和普遍性。我们称$F^{(n)}$为n维空间V（在基底ε_1，ε_2，…，ε_n上）的坐标

空间。

我们知道，如果V是n维空间，一般来说，V的基底是很多的。我们需要了解各个基底之间的一般联系。

命题1 若 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是V的一个基底，$\xi_1, \xi_2, \cdots, \xi_n$是$V$中$n$个向量。令

$$(\xi_1\xi_2\cdots\xi_n) = (\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)P, \tag{7}$$

那么 $\xi_1, \xi_2, \cdots, \xi_n$ 是V的基底当且仅当P是可逆的，即$\det P \neq 0$。

证明 充分性是简单的。因为P是可逆的，于是必有P^{-1}。用P^{-1}从右侧去乘（7）式两端，便得

$$(\xi_1\xi_2\cdots\xi_n)P^{-1} = ((\varepsilon_1\varepsilon_2\cdots\varepsilon_n)P)P^{-1} = (\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)$$

这表明 $\xi_1, \xi_2, \cdots, \xi_n$ 与$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 等价。所以$\xi_1, \xi_2, \cdots, \xi_n$ 也是V的一个基底。

反之，如果如（7）式表出的 $\xi_1, \xi_2, \cdots, \xi_n$ 是V的一个基底，从而也有表出式

$$(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n) = (\xi_1\ \xi_2\ \cdots\xi_n)Q \tag{8}$$

把（7）式代入（8）中，则得

$$(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n) = (\varepsilon_1\ \varepsilon_2\cdots\varepsilon_n)PQ \tag{9}$$

由于 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是基底，从而线性表出式（9）中的诸系数PQ是唯一的，所以必有 $PQ=I$。这就证明P是可逆的。

我们把联结两个基底的关系式（7）叫做从基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到基底 $\xi_1, \xi_2, \cdots, \xi_n$ 的变换公式。当中的可逆阵P叫做过渡阵。

命题2 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 与 $\delta_1, \delta_2, \cdots, \delta_n$ 是V的两个基底，向量ξ在这两个基底上的坐标分别是

$$x_1,\ x_2,\ \cdots,\ x_n \text{ 与 } y_1,\ y_2,\ \cdots,\ y_n$$

如果从 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\delta_1, \delta_2, \cdots, \delta_n$ 的过渡阵为P，那么

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = P^{-1}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{10}$$

证明 由ξ在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 上的坐标为 $x_1, x_2, \cdots, x_n$ 即有

$$\xi = (\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

同理

$$\xi = (\delta_1\ \delta_2\ \cdots\ \delta_n)\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

利用从 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\delta_1, \delta_2, \cdots, \delta_n$ 的过渡关系，则得

$$\xi = (\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\ P\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

由ξ在基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 上的坐标是唯一的，所以

$$P\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}，\ 即\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = P^{-1}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

(10) 式叫做**坐标变换公式**。

例1 在 $F_4[x]$ 中$1, x, x^2, x^3$ 是一个基底，于是任一向量

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

在这个基底上的坐标应当是 a_0, a_1, a_2, a_3。如果具体给出一组向量

$$f_1(x) = 1 + x + x^2$$

$$f_2(x) = 1 + 2x + 3x^2$$

$$f_3(x) = -1 + x^2$$

那么 $f_1(x), f_2(x), f_3(x)$ 在基底 $1, x, x^2, x^3$ 上的坐标为

$$\overline{f_1(x)} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix},\ \overline{f_2(x)} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix},\ \overline{f_3(x)} = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

从而向量组 $f_1(x)$, $f_2(x)$, $f_3(x)$ 的坐标列矩阵为

$$A=\begin{pmatrix}1 & 1 & -1\\ 1 & 2 & 0\\ 1 & 3 & 1\\ 0 & 0 & 0\end{pmatrix}.$$

容易计算 $\operatorname{rank} A=2$，所以 $f_1(x)$, $f_2(x)$, $f_3(x)$ 是 $F_4[x]$ 中一组线性相关向量，并且由于 A 的前两列上有基础子式，因而 $f_1(x)$, $f_2(x)$ 是 $f_1(x)$, $f_2(x)$, $f_3(x)$ 的一个极大线性无关组.

例 2 在 $F^{(3)}$ 中 $\varepsilon_1=(1, 0, 0)$, $\varepsilon_2=(0, 1, 0)$, $\varepsilon_3=(0, 0, 1)$ 是一个基底. 对任一向量 $\xi=(x, y, z)$, 由

$$\xi=x\varepsilon_1+y\varepsilon_2+z\varepsilon_3$$

可知 ξ 在基底 ε_1, ε_2, ε_3 上的坐标就是 x, y, z.

若令 $\delta_1=\varepsilon_1+\varepsilon_2$, $\delta_2=\varepsilon_2+\varepsilon_3$, $\delta_3=\varepsilon_1+\varepsilon_3$, 则有

$$(\delta_1\ \delta_2\ \delta_3)=(\varepsilon_1\ \varepsilon_2\ \varepsilon_3)\begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix}.$$

由于

$$P=\begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix}, \quad \det P=2$$

所以 P 是可逆的，故而 δ_1, δ_2, δ_3 也是 $F^{(3)}$ 的一个基底，从 ε_1, ε_2, ε_3 到 δ_1, δ_2, δ_3 的过渡阵就是 P.

在这个新基底 δ_1, δ_2, δ_3 上向量 $\xi=(x, y, z)$ 也有坐标，设为 x', y', z', 即有

$$\xi=x'\delta_1+y'\delta_2+z'\delta_3.$$

由坐标变换公式，可计算出

$$\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}=P^{-1}\begin{pmatrix}x\\ y\\ z\end{pmatrix}, \quad \text{其中}\quad P^{-1}=\begin{pmatrix}\frac{1}{2} & \frac{1}{2} & -\frac{1}{2}\\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2}\end{pmatrix}.$$

具体地，如 $\xi=(1, 2, 3)$，那么按这个坐标变换公式，$\xi=(1, 2, 3)$ 在 $\delta_1, \delta_2, \delta_3$ 上的坐标为

$$x'=\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 2+\left(-\frac{1}{2}\right)\cdot 3=0,$$

$$y'=-\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 2+\frac{1}{2}\cdot 3=2,$$

$$z'=\frac{1}{2}\cdot 1+\left(-\frac{1}{2}\right)\cdot 2+\frac{1}{2}\cdot 3=1.$$

即 $(1, 2, 3)=2\delta_2+\delta_3$.

练　习　六

1. 同一向量在不同基底上的坐标是否一定不同？不同向量在同一基底上的坐标是否可以相同？

2. 在 $F_{n+1}[x]$ 中，试求两组基底，使向量

$$g(x)=(x+a)^n$$

在这两组基底上的坐标相同.

3. 证明 x^2+x, x^2-x, $x+1$ 是 $F_3[x]$ 的基底，并求 $2x^2+7x+3$ 在此基底上的坐标，

4. 试给出 $M_2(F)$ 中两组不同的基底，并求向量

$$A=\begin{pmatrix}-1 & 3\\ 0 & 2\end{pmatrix}.$$

在所给的基底上的坐标.

5. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性空间 V 的基底，求由它到基底 $\alpha_2, \alpha_3, \cdots, \alpha_n, \alpha_1$ 的过渡阵.

6. 在 $D^{(3)}$ 中，1）求由基底

$$\varepsilon_1=(1, 0, 0),\quad \varepsilon_2=(0, 1, 0),\quad \varepsilon_3=(0, 0, 1)$$

到基底

$$\eta_1=(1, 0, 0),\ \eta_2=(1, 1, 0),\ \eta_3=(1, 1, 1)$$

的过渡阵。

2）求由基底 $\{\eta_i\}$，经过渡阵

$$A=\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

所得到的新基底。

7．在 $F_4[x]$ 中，1）求由基底

$$1,\ x,\ x^2,\ x^3 \tag{1}$$

到基底

$$1,\ 1+x,\ 1+x+x^2,\ 1+x+x^2+x^3 \tag{2}$$

的过渡阵。

2）已知 $g(x)$ 在基底（1）上的坐标为1，0，－2，5；$f(x)$ 在基底（2）上的坐标为7，0，8，－2；求 $f(x)+g(x)$分别在基底（1），（2）上的坐标。

§7 线性空间的同构

在上节我们已经看到，对任一有限维线性空间V，只要它的维数n确定了，V就完全的具体的转化为 n 数组空间 $F^{(n)}$。二者的空间结构（由线性运算所决定的性质）就完全是一样的。如果说维数相同的各个线性空间有什么差异的话，那也仅仅是它们的向量特定的表现形式有所不同。而这种向量的表现形式的不同对空间的结构来说并没有实质性的影响。归根结底一句话，凡n维线性空间，其向量都由n数组所唯一决定，并且向量的线性运算也是由这些n数组的线性运算来实现，即凡n维线性空间都可由n数组空间来实现。为了完整的确切的表达这种观念和方法，使我们引出代数学的一个基本概念——同构。

定义 设V_1，V_2 都是F上的线性空间。如果在V_1与V_2之间存在一个一一对应φ，使得若

$$\varphi: \alpha \longmapsto \alpha'$$
$$\beta \longmapsto \beta'$$

那么

$$\alpha+\beta \longmapsto \alpha'+\beta',$$
$$k\alpha \longmapsto k\alpha'.$$

则称φ是V_1与V_2的一个同构映射，并说V_1与V_2是同构的，记作$V_1 \simeq V_2$。

例 1 考虑 $F_3[x]$ 与 $F^{(3)}$ 这两个线性空间。容易想到它们之间存在一个自然的一一对应，即

$$\varphi: f(x)=a_0+a_1x+a_2x^2 \longmapsto (a_0, a_1, a_2).$$

已经知道，这个映射φ刚好满足同构映射的两条要求，即φ是 $F_3[x]$ 与$F^{(3)}$ 的同构映射，$F_3[x] \simeq F^{(3)}$。

例 2 考虑 $M_2(F)$ 与 $F^{(4)}$，它们之间的映射

$$\varphi: \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto (a, b, c, d).$$

显然，φ 是一个一一对应，并且满足同构映射的两条要求，所以 $M_2(F) \simeq F^{(4)}$。

例 3 若 $\dim V=n$ 则 $V \simeq F^{(n)}$。

事实上 任取V的一个基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$。于是映射

$$\varphi: \alpha \longmapsto \bar{\alpha}, \ \bar{\alpha}\text{是 } \alpha \text{ 在 } \varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n \text{ 上的坐标列},$$

就是n维空间V与 $F^{(n)}$ 的同构映射，$V \simeq F^{(n)}$。

例 4 如果 $V_1 \simeq V_2$，那么 $V_2 \simeq V_1$。

事实上 由 $V_1 \simeq V_2$，可设φ是V_1 与 V_2 的一个同构映射。考虑φ的逆映射ψ，当然ψ就是从V_2到V_1的一个一一对应：

$$\psi: \alpha' \longmapsto \alpha, \text{ 其中 } \varphi: \alpha \longmapsto \alpha', \ \forall \alpha' \in V_2.$$

于是若

$$\psi: \beta' \longmapsto \beta, \text{ 其中 } \varphi: \beta \longmapsto \beta'.$$

那么由

$$\varphi(\alpha+\beta)=\varphi(\alpha)+\varphi(\beta)=\alpha'+\beta'$$

$$\varphi(k\alpha)=k\varphi(\alpha)=k\alpha'$$

可知

$$\psi: \ \alpha'+\beta' \longmapsto \alpha+\beta$$
$$k\alpha' \longmapsto k\alpha$$

所以ψ是V_2与V_1的同构映射，$V_2 \simeq V_1$。

例 5　如果φ是V_1与V_2的同构映射，那么 $\varphi(0)=0$（注意：等号前边的 0 是V_1的零向量，等号后边的 0 则是V_2的零向量）。

事实上　$\varphi(0)=\varphi(k0)=k\varphi(0)$。令 $k=0$时，便得 $\varphi(0)=0$。

例 6　如果φ是V_1与V_2的同构映射，那么对于V_1的任意一组向量 α_1，α_2，$\cdots$，α_s，都有

$$\varphi(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=k_1\varphi(\alpha_1)+k_2\varphi(\alpha_2)+ +\cdots+k_s\varphi(\alpha_s).$$

事实上　当 $s=2$ 时，则有

$$\varphi(k_1\alpha_1+k_2\alpha_2)=\varphi(k_1\alpha_1)+\varphi(k_2\alpha_2) =k\varphi(\alpha_1)+k_2\varphi(\alpha_2)$$

归纳假设对 $s-1$ 个向量等式成立，看s个向量的情形，于是

$$\varphi(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=\varphi((k_1\alpha_1+k_2\alpha_2+\cdots+k_{s-1}\alpha_{s-1})+k_s\alpha_s)$$
$$=\varphi(k_1\alpha_1+k_2\alpha_2+\cdots+k_{s-1}\alpha_{s-1})+\varphi(k_s\alpha_s)$$
$$=k_1\varphi(\alpha_1)+k_2\varphi(\alpha_2)+\cdots+k_{s-1}\varphi(\alpha_{s-1})+k_s\varphi(\alpha_s).$$

命题 1　**设φ是V_1与V_2的一个同构映射**，α_1，α_2，$\cdots$，$\alpha_s \in V_1$，$\varphi(\alpha_i)=\alpha'_i$（$i=1$，$2$，$\cdots$，$s$）**那么**

$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$ **当且仅当**$k_1\alpha_1'+k_2\alpha'_2+\cdots+k_s\alpha_s'=0$。

证明　对任意的 k_1，k_2，$\cdots$，k_s 都有

$$\varphi(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=k_1\alpha_1'+k_2\alpha_2'+\cdots+k_s\alpha_s'$$

而在同构映射之下，一个向量 ξ 的象是零向量当且仅当 $\xi=0$。由此便得要证明的结果。

推论　α_1，α_2，$\cdots$，α_s 是线性相关的当且仅当α_1'，α_2'，$\cdots$，α_s'是线性相关的。

命题 2　**设 V_1 与V_2都是F上有限维空间。那么 $V_1 \simeq V_2$ 当且**

仅当 $\dim V_1=\dim V_2$。

证明 如果 $V_1\simeq V_2$，由命题1可知 $\dim V_1=\dim V_2$，反之，如果 $\dim V_1=\dim V_2$。令 $\dim V_1=n$。于是 $V_1=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$，$V_2=[\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n']$。这样可以自然的规定$V_1$到$V_2$的一个映射

$$\varphi: \alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n \longmapsto \alpha'=a_1\varepsilon_1'+a_2\varepsilon_2'+\cdots+a_n\varepsilon_n'.$$

不难知道，映射φ是V_1与V_2的同构映射，$V_1\simeq V_2$。

命题3 令X表示数域F上一切线性空间的集合。于是同构关系$\simeq$是X的一个等价关系。

证明 对任意的 V_1，V_2，$V_3\in X$，都有

1）反身性 $V_1\simeq V_1$

这只要令

$$\varphi: \alpha\longmapsto\alpha, \quad \forall\alpha\in V_1$$

便得 $V_1\simeq V_1$。

2）对称性 若 $V_1\simeq V_2$，那么令φ是V_1与V_2的同构映射时，φ的逆映射ψ就是V_2与V_1的同构映射，即 $V_2\simeq V_1$。

3）传递性 若 $V_1\simeq V_2$，$V_2\simeq V_3$。令 φ_1，φ_2 分别是相应的同构映射，那么 $\varphi_2\varphi_1$ 就是V_1与V_3的同构映射，即有 $V_1\simeq V_3$。

同构既然是X的等价关系，那么同构关系给X决定一个分类Σ，即

V_1与V_2分在一类里当且仅当 $V_1\simeq V_2$。

我们已经知道，与有限维空间同构的必定是有限维空间，并且两个有限维空间同构的充分必要条件是二者同维数。这样，X的分类Σ大体上可表达为

$$\Sigma=\{[\{0\}], [F^{(1)}], [F^{(2)}], \cdots, [F^{(n)}], \cdots, [L], \cdots\}.$$

其中 $[F^{(r)}]$ 表示与 $F^{(r)}$ 同构的空间组成的类，$[L]$及其以后的都是无限维空间组成的类，当然$[\{0\}]$就是与零空间$\{0\}$同构的空间组成的类。

练 习 七

1．证明：复数域上的二阶方阵的集合 $M_2(C)$ 做为实数域 D 上的线性空间与 $D^{(8)}$ 同构。

2．证明 D^+ 按练习二中第2题规定的运算所构成的线性空间与 D 看作自身上的线性空间同构。

3．证明 $F[x]$ 与它的真子空间同构。

习 题 十 一

1．设 $M=\{(a,\ b)\mid a,\ b\in D\}$，规定

$$(a_1,\ b_1)\oplus(a_2,\ b_2)=(a_1+a_2,\ b_1+b_2+a_1a_2)$$

$$k(a_1,\ b_1)=(ka_1,\ kb_1+\frac{1}{2}k(k-1)a_1^2)$$

证明 M 是实数域 D 上的线性空间。

2．平面上全体向量构成的集合 π，对于通常的向量加法和如下规定的倍数乘法

$$k\alpha=\alpha$$

$k\in D$，问 π 是否是实数域 D 上的线性空间。

3．设 W_1，W_2 是数域 F 上的线性空间 V 的子空间。令 α,β 是 V 中两个向量，如果 $\alpha\in W_2$，但 $\alpha\overline{\in}W_1$，$\beta\overline{\in}W_2$，证明

1）对任意 $k\in F$，都有 $\beta+k\alpha\overline{\in}W_2$

2）至多有一个 $k\in F$，使得 $\beta+k\alpha\in W_1$

4．设 W_1，W_2 是线性空间 V 的真子空间，证明 V 中存在一个向量 α 使得 $\alpha\overline{\in}W_1$，同时 $\alpha\overline{\in}W_2$。

5．设 W_1，W_2，…，W_r 是线性空间 V 的真子空间，证明 V 中存在一个向量 α 同时不属于 W_i $(i=1,\ 2,\ \cdots,\ r)$。

6．在有理数域Q上，实数集D所构成的线性空间中，求由1，$\sqrt{2}$，$\sqrt{5}$生成的子空间的维数。

7．设W_1，W_2是n维线性空间V的子空间，如果

$$\dim(W_1+W_2)=\dim(W_1\cap W_2)+1 \tag{1}$$

证明和W_1+W_2必与W_1，W_2中某一个重合，而交$W_1\cap W_2$必与另一个重和。

8．在$F_{n+1}[x]$中，如果$f(x)$是n次多项式，证明

$$f(x),\ f'(x),\ \cdots,\ f^{(n)}(x)$$

是$F_{n+1}[x]$的基底。如果

$$f(x)=(x-a)^n$$

求对任意$g(x)\in F_{n+1}[x]$的坐标。

9．设α_1，α_2，…，α_n是线性空间$V_n(F)$的基底，

1）若$k_i\in F$，$(i=1,\ 2,\ \cdots,\ n)$不全为0，证明以方程

$$k_1x_1+k_2x_2+\cdots+k_nx_n=0$$

的解c_1，c_2，…，c_n为坐标的向量，即

$$c_1\alpha_1+c_2\alpha_2+\cdots+c_n\alpha_n$$

的集合L_1是$V_n(F)$的$n-1$维子空间。

2）证明以方程组

$$\begin{aligned}
&a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0\\
&a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0\\
&\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\quad\cdots\\
&a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0
\end{aligned}$$

的解c_1，c_2，…，c_n为坐标的向量集合L_2是$V_n(F)$的子空间，并求其维数。

10．证明$D^{(n)}$的任意一个子空间都是某个齐次线性方程组的解空间。

11．证明$D^{(n)}$的任意真子空间都是若干个$n-1$维子空间的交。

12．设α，β，$\gamma\in V$，如果

$$c_1\alpha+c_2\beta+c_3\gamma=0$$

当 $c_1 \cdot c_3 \neq 0$ 时，证明

$$L(\alpha,\ \beta) = L(\beta,\ \gamma)$$

13．设W_1，W_2 分别是齐次线性方程组

$$x_1 + x_2 + \cdots + x_n = 0 \tag{1}$$

与

$$x_1 = x_2 = \cdots = x_n \tag{2}$$

的解空间，证明：

$$D^{(n)} = W_1 \dotplus W_2$$

14．设W是线性空间 $V_n(F)$ 的非平凡子空间 $W \neq V_n(F)$，$W \neq \{0\}$，证明W 不只有一个补空间。

15．设 $M_n(F)$ 表示数域F上的一切n阶方阵所构成的线性空间，令

$$S = \{A \in M_n(F) \mid A' = A\}$$

$$T = \{A \in M_n(F) \mid A' = -A\}$$

证明：$M_n(F) = S \dotplus T$。

第十二章　线性变换

与线性空间理论紧密相关的是线性变换的理论．它一方面可以看做是最简单而又最基本的线性函数的一般化。另一方面是几何空间中一类相当广泛的变换在一般线性空间上的推广．特别地，对有限维线性空间来说，线性变换理论与矩阵的相似理论是完全相当的．本章讨论两方面的问题：线性变换的一般概念与基本性质；有限维线性空间上线性变换的标准形．

§1　线性变换的定义及简单性质

为了叙述自然，使读者易于接受，我们从一些熟悉的例子入手，从而导出线性变换的一般概念．

例1　设D为实数域，大家知道，最简单而又最基本的函数莫过于线性函数

$$y=f(x)=ax \tag{1}$$

它的定义域和值域都是实数域D，即对D中每一个实数x，线性函数"f"使其对应一个函数值ax，并且具有性质

1) $f(x_1+x_2)=f(x_1)+f(x_2)$，

2) $f(kx)=kf(x)$．

用话来说就是，x_1与x_2的和x_1+x_2的函数值等于x_1与x_2的函数值$f(x_1)$与$f(x_2)$的和$f(x_1)+f(x_2)$，x的k倍kx的函数值等于x的函数值$f(x)$的k倍$kf(x)$．

我们这里是在考虑线性空间上的线性变换问题．因此，最好用线性空间的语言把上述事实表达出来，以便从中得出规律性的认

识。

大家知道，任一数域F都自然可以看成是F上的线性空间。因而其中每一个数x都叫向量。这样，定义在实数域D上的线性函数f：

$$f: x \longmapsto ax \quad 或 \ f(x)=ax$$

就可以看成是线性空间D上的一个变换f。并且线性函数的两条基本性质便可相应的说成：

1）两个向量x_1与x_2的和x_1+x_2的象$f(x_1+x_2)$等于象的和$f(x_1)+f(x_2)$；

2）向量x的k倍kx的象$f(kx)$等于象的k倍$kf(x)$。

使用这样的语言就能对以往所熟悉的一些比较分散的对象得到统一的认识，从而掌握它们共同的规律性。

例2　考虑几何空间$D^{(2)}$。它可以看成是有了笛卡儿直角坐标系的一张平面，$D^{(2)}$的变换

$$\sigma: \alpha=(x, y) \longmapsto (x, 0)$$

具有与线性函数同样的两条性质

$$1)\ \sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),$$

$$2)\ \sigma(k\alpha)=k\sigma(\alpha), \quad \forall \alpha, \beta\in D^{(2)}, k\in D.$$

这里的σ就是平面$D^{(2)}$的往x轴上的“射影变换”，σ把向量$\alpha=(x, y)$射成α在x轴上的“射影”$(x, 0)$。

例3　考虑$M_n(F)$，这是F上n^2维的线性空间，它的转置变换

$$\tau: A \longmapsto A', \quad \forall A\in M_n(F),$$

具有与线性函数同样的两条性质：

$$1)\ \tau(A+B)=\tau(A)+\tau(B),$$

$$2)\ \tau(kA)=k\tau(A).$$

综合以上三例，我们便可导出线性变换的一般概念。

定义　设V是F上任一线性空间。σ是V的一个变换，如果

$$1)\ \sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta),$$

2) $\sigma(k\alpha)=k\sigma(\alpha)$，$\forall\alpha$，$\beta\in V$，$k\in F$，

则称σ是V的一个线性变换。

以下常用希腊字母σ，τ，ρ，… 等表示线性变换。

按此定义，线性函数 $f(x)=ax$ 是线性空间D上的线性变换；射影变换$\sigma((x, y))=(x, 0)$ 是平面$D^{(2)}$ 上的线性变换；转置变换$\tau(A)=A'$ 是线性空间 $M_n(F)$ 上的线性变换。

例 4 考虑一元多项式的线性空间$D[x]$，求导数的方法

$$d:\ f(x)\longmapsto f'(x)$$

是$D[x]$ 的一个变换，早已知道，求导数的两条性质：

1) $d(f(x)+g(x))=d(f(x))+d(g(x))$，

2) $d(kf(x))=kd(f(x))$。

这样，求导变换d是线性空间 $D[x]$ 的一个线性变换，同理可知，求积变换

$$s:\ f(x)\longmapsto\int_0^x f(t)\,dt$$

也是$D[x]$ 的一个线性变换。

例 5 F上的任一n维空间$V=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$。取定一个$n$阶方阵$A\in M_n(F)$。这样，对$V$中任意的向量 $\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$，我们规定

$$\sigma:\ \alpha\longrightarrow\alpha',\qquad \alpha'=a_1'\varepsilon_1+a_2'\varepsilon_2+\cdots+a_n'\varepsilon_n$$

这里

$$A\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}=\begin{pmatrix}a_1'\\a_2'\\\vdots\\a_n'\end{pmatrix}.$$

直接用坐标来写，就是

$$\sigma:\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\longmapsto\begin{pmatrix}a_1'\\a_2'\\\vdots\\a_n'\end{pmatrix}.$$

显然，利用矩阵A如此规定的方法σ是n维线性空间V的一个变换。

进而也不难验证，σ 是线性的. 如果 $\beta = b_1\varepsilon_1 + b_2\varepsilon_2 + \cdots + b_n\varepsilon_n$，那么

$$\sigma:\ \beta \longrightarrow \beta', \qquad \beta' = b_1'\varepsilon_1 + b_2'\varepsilon_2 + \cdots + b_n'\varepsilon_n$$

其中

$$A\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} b_1' \\ b_2' \\ \vdots \\ b_n' \end{pmatrix}.$$

于是 $\alpha + \beta = (a_1 + b_1)\varepsilon_1 + (a_2 + b_2)\varepsilon_2 + \cdots + (a_n + b_n)\varepsilon_n$.

$$A\begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix} = A\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + A\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1' \\ a_2' \\ \vdots \\ a_n' \end{pmatrix} + \begin{pmatrix} b_1' \\ b_2' \\ \vdots \\ b_n' \end{pmatrix},$$

$k\alpha = (ka_1)\varepsilon_1 + (ka_2)\varepsilon_2 + \cdots + (ka_n)\varepsilon_n$,

$$A\begin{pmatrix} ka_1 \\ ka_2 \\ \vdots \\ ka_n \end{pmatrix} = k\cdot A\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = k\begin{pmatrix} a_1' \\ a_2' \\ \vdots \\ a_n' \end{pmatrix}.$$

这就证明了：

$$\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta),$$

$$\sigma(k\alpha) = k\sigma(\alpha).$$

即σ是V的线性变换.

例 6　设V是任一线性空间. 考虑V的如下的变换：

$$\sigma_1:\ \alpha \longmapsto \alpha,$$

$$\sigma_2:\ \alpha \longmapsto 0.$$

这里σ_1就是把向量α射成α本身，即使V中每一个向量α都不动的变换，而σ_2则是把每一个向量α都射成零向量. 显然，这两个很特殊的变换都是线性变换，σ_1叫做**恒等变换**，特别地，记作 ε，即 $\varepsilon(\alpha) = \alpha$，$\forall \alpha \in V$，$\sigma_2$叫做**零变换**，特别地，记作 0，即$0(\alpha) = 0$，$\forall \alpha \in V$.

命题　设 σ 是线性空间V的一个变换. σ 是线性变换当且仅当 $\sigma(k_1\alpha_1 + k_2\alpha_2) = k_1\sigma(\alpha_1) + k_2\sigma(\alpha_2)$，$\forall \alpha_1, \alpha_2 \in V$，$k_1, k_2 \in F$.

证明　必要性是明显的，反过来，令 $k_1=1$，$k_2=1$ 就得出线性变换所需要的第一个条件，令 $k_2=0$，就得出线性变换的第二个条件。

推论 1　**若σ是线性变换，则有**

$$\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_s\sigma(\alpha_s),$$

特别地，$\sigma(0)=0$，$\sigma(-\alpha)=-\sigma(\alpha)$。

对s做数学归纳法就可证明所说的等式。

推论 2　**若 σ 是线性变换，** $\alpha_1,\alpha_2,\cdots,\alpha_s$ **是一组线性相关的向量，那么它们的象** $\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_s)$ **也一定是线性相关的。换言之，线性变换σ把线性相关的向量还变成线性相关的向量。**

事实上　由 α_1，α_2，$\cdots$，α_s是线性相关的，则有不全为零的数 k_1，k_2，$\cdots$，k_s 使

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0.$$

于是

$$\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=\sigma(0)=0.$$

即

$$k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_s\sigma(\alpha_s)=0.$$

所以 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，$\cdots$，$\sigma(\alpha_s)$ 是线性相关的。

由推论 2 可知：

$\{\sigma(\alpha_1),\ \sigma(\alpha_2),\ \cdots,\ \sigma(\alpha_s)\}$ 的秩 $\leqslant\{\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_s\}$ 的秩。

练　习　一

1. 设σ_i是$D^{(2)}$ 中的一个变换，对$D^{(2)}$ 中任意向量 $\alpha=(x_1,x_2)$ 有

1) $\sigma_1(\alpha)=(x_1,\ -x_2)$

2) $\sigma_2(\alpha)=(-x_1,\ x_2)$

3) $\sigma_3(\alpha)=(-x_1,\ -x_2)$

4) $\sigma_4(\alpha)=(x_1, 0)$

5) $\sigma_5(\alpha)=(0, x_2)$

证明 σ_1, σ_2, σ_3, σ_4, σ_5 是 $D^{(2)}$ 的线性变换．并说明几何意义．

2．证明$D^{(2)}$ 中的旋转变换，即对任意 $\alpha=(x_1, x_2)$属于$D^{(2)}$有

$$\sigma(x_1, x_2)=(x_1\cos\theta-x_2\sin\theta, x_1\sin\theta+x_2\cos\theta)$$

是线性变换．而平移变换，即

$\tau(x_1, x_2)=(x_1+a, x_2+b)$，$a$，$b$ 不全为零，不是线性变换．

3．在$D^{(3)}$ 中，对任意向量 $\alpha=(x_1, x_2, x_3)\in D^{(3)}$，下列变换哪些是线性变换．

1) $\sigma_1(\alpha)=(x_1+x_2, x_2+x_3, x_3+x_1)$

2) $\sigma_2(\alpha)=(x_1^2, x_2^2, x_3^2)$

3) $\sigma_3(\alpha)=(\cos x_1, \sin x_2, 0)$

4．设 $M_n(F)$ 是数域F上的一切n阶方阵所构成的线性空间，A，B是 $M_n(F)$ 中固定的n阶方阵，X 是 $M_n(F)$ 中任意矩阵，证明变换

$$\sigma(X)=AXB$$

是线性变换．

5．证明数域F上的一维线性空间V的一个变换σ是线性变换的充分必要条件是对V中任意向量α都有

$$\sigma(\alpha)=\lambda\alpha.$$

6．设 $A\in M_n(F)$，为一固定向量，对任意的 $X\in M_n(F)$，定义变换

$$\sigma(X)=AX-XA$$

1）证明σ是线性变换．

2）对任意X，$Y\in M_n(F)$ 有

$$\sigma(XY)=(\sigma(X))Y+X(\sigma(Y))$$

§2　线性变换的运算

我们从线性函数的运算说起．设 $f(x)=ax$，$g(x)=bx$ 是两个线性函数（自然可以看做是线性空间 D 上的两个线性变换）．我们已知两个函数的加法就是使函数值相加，从而得到一个函数．比如

$$h(x)=f(x)+g(x)=ax+bx=(a+b)x.$$

这个 $h(x)$ 叫做 $f(x)$ 与 $g(x)$ 的和，记作 $h=f+g$．再有，一个函数与一个数k的倍数运算，就是用数k去乘函数值，从而得到一个函数．比如

$$u(x)=k\cdot f(x)=k\cdot ax=(ka)x.$$

称 $u(x)$ 为k与 $f(x)$ 的积，记作 $u=kf$．

最后，两个线性函数可以复合而成一个函数．比如

$$v(x)=f(g(x))=f(bx)=a(bx)=(ab)x.$$

以后称 $v(x)$ 为 $f(x)$ 与 $g(x)$ 的积．记作 $v=fg$．

总起来说，关于线性函数有三种运算方法：

加法　$(f+g)(x)=f(x)+g(x)$，

倍数　$(kf)(x)=k\cdot f(x)$，

乘法　$(fg)(x)=f(g(x))$．

对于一般的线性变换也有完全类似的三种运算．现在分别规定如下．

设V是F上任一线性空间．$L(V)$ 表示V 上一切线性变换组成的集合．我们给 $L(V)$ 规定三种运算．

加法　$\forall\sigma$，$\tau\in L(V)$，令

$$(\sigma+\tau)(\alpha)=\sigma(\alpha)+\tau(\alpha),\quad\forall\alpha\in V.$$

倍数　$\forall k\in F$，$\sigma\in L(V)$，令

$$(k\sigma)(\alpha)=k\cdot\sigma(\alpha),\quad\forall\alpha\in V,$$

乘法　$\forall\sigma$，$\tau\in L(V)$，令

$$(\sigma\tau)(\alpha)=\sigma(\tau(\alpha)),\quad\forall\alpha\in V.$$

显然，如此得到的 $\sigma+\tau$，$k\sigma$，$\sigma\tau$ 都是线性空间V的变换．下面我们证明它们还都是V的线性变换．

命题1 设 σ，τ 是线性空间 V 的线性变换，$k\in F$．那么 $\sigma+\tau$，$k\sigma$，$\sigma\tau$ 都是V的线性变换．

证明 $\forall\alpha$，$\beta\in V$，k_1，$k_2\in F$，都有

$$\begin{aligned}(\sigma+\tau)(k_1\alpha+k_2\beta)&=\sigma(k_1\alpha+k_2\beta)+\tau(k_1\alpha+k_2\beta)\\&=k_1\sigma(\alpha)+k_2\sigma(\beta)+k_1\tau(\alpha)+k_1\tau(\beta)\\&=k_1\sigma(\alpha)+k_1\tau(\alpha)+k_2\sigma(\beta)+k_2\tau(\beta)\\&=k_1(\sigma(\alpha)+\tau(\alpha))+k_2(\sigma(\beta)+\tau(\beta))\\&=k_1(\sigma+\tau)(\sigma)+k_2(\sigma+\tau)(\beta)\end{aligned}$$

由§1命题，$\sigma+\tau$ 是线性的．

同样可证 $k\sigma$，$\sigma\tau$ 都是线性的．

例1 在$D^{(3)}$中，令

$$\sigma:\ (x,\ y,\ z)\longmapsto(x,-y,z),$$

$$\tau:\ (x,\ y,\ z)\longmapsto(-x,y,z)\qquad\forall\xi=(x,y,z)\in D^{(3)}.$$

显然σ与τ都是$D^{(3)}$的线性变换，于是

$$\begin{aligned}(\sigma+\tau)(\xi)&=\sigma(\xi)+\tau(\xi)\\&=(x,\ -y,\ z)+(-x,\ y,\ z)\\&=(0,\ 0,\ 2z),\end{aligned}$$

$$\begin{aligned}(k\sigma)(\xi)&=k\cdot\sigma(\xi)\\&=k(x,\ -y,\ z)\\&=(kx,\ -ky,\ kz),\end{aligned}$$

$$\begin{aligned}(\sigma\tau)(\xi)&=\sigma(\tau(\xi))\\&=\sigma((-x,\ y,\ z))\\&=(-x,\ -y,\ z).\end{aligned}$$

命题2 线性变换的运算具有以下性质：

1) $\sigma+\tau=\tau+\sigma$

2) $(\sigma+\tau)+\rho=\sigma+(\tau+\rho)$

3) $0+\sigma=\sigma=\sigma+0$

4）若规定$(-\sigma)(\alpha)=-\sigma(\alpha)$，称$-\sigma$为$\sigma$的负变换，那么$-\sigma$是线性变换，并且

$$\sigma+(-\sigma)=0=(-\sigma)+\sigma.$$

5）$k(\sigma+\tau)=k\sigma+k\tau$

6）$(l+k)\sigma=l\sigma+k\sigma$

7）$(lk)\sigma=l(k\sigma)$

8）$1\sigma=\sigma$

9）$(\sigma\tau)\rho=\sigma(\tau\rho)$

10）$\varepsilon\sigma=\sigma=\sigma\varepsilon$

11）$\sigma(\tau+\rho)=\sigma\tau+\sigma\rho$

$(\tau+\rho)\sigma=\tau\sigma+\rho\sigma$

12）$k(\sigma\tau)=(k\sigma)\tau=\sigma(k\tau)$.

证明　这12条性质都是容易由定义直接证明的。我们只证明11）与12）两条。余者留给读者去验证。

先证11）中的左分配律成立。$\forall\alpha\in V$，于是

$$\begin{aligned}\sigma(\tau+\rho)(\alpha)&=\sigma[(\tau+\rho)(\alpha)]\\&=\sigma(\tau(\alpha)+\rho(\alpha))\\&=\sigma(\tau(\alpha))+\sigma(\rho(\alpha))\\&=\sigma\tau(\alpha)+\sigma\rho(\alpha)\\&=(\sigma\tau+\sigma\rho)(\alpha)\end{aligned}$$

所以 $\sigma(\tau+\rho)=\sigma\tau+\sigma\rho$.

同样可证右分配律成立。

再证12）成立

$$\begin{aligned}(k\sigma)\tau(\alpha)&=k\sigma(\tau(\alpha))\\&=k(\sigma(\tau(\alpha)))\\&=k(\sigma\tau(\alpha))\\&=k(\sigma\tau)(\alpha)\end{aligned}$$

所以 $(k\sigma)\tau=k(\sigma\tau)$。同样可证 $\sigma(k\tau)=k(\sigma\tau)$。

以上命题2中的前八条性质说明，线性空间V的一切线性变换

所组成的集合 $L(V)$，用线性变换的加法与倍数运算，做成了线性空间。对线性空间 $L(V)$ 来说，它的向量就是线性空间V的线性变换。

由于存在一个特殊的线性变换—恒等变换 ε，它的特性是对任一线性变换 σ 都有

$$\varepsilon\sigma = \sigma = \sigma\varepsilon,$$

这样，在一切线性变换的集合 $L(V)$ 中， 应当考虑可逆变换的问题。

设 $\sigma \in L(V)$。如果有 $\tau \in L(V)$，使

$$\sigma\tau = \varepsilon = \tau\sigma$$

则称σ是可逆线性变换，简称σ是可逆的。

命题 3 σ 是可逆的当且仅当σ有逆变换。

证明 我们分以下四点来说明。

1）如果线性变换σ有逆变换，记作σ^{-1}，那么 σ^{-1} 也是线性变换。

事实上 令

$$\sigma^{-1}: \alpha' \longmapsto \alpha, \quad \beta' \longmapsto \beta.$$

则有 $\sigma(\alpha) = \alpha'$，$\sigma(\beta) = \beta'$. 从而

$$\sigma: \alpha + \beta \longmapsto \alpha' + \beta'$$
$$k\alpha \longmapsto k\alpha'$$

这就表明 $\sigma^{-1}(\alpha' + \beta') = \sigma^{-1}(\alpha') + \sigma^{-1}(\beta')$， $\sigma^{-1}(k\alpha') = k\sigma^{-1}(\alpha')$，即 σ^{-1} 是线性的。

2）$\sigma\sigma^{-1} = \varepsilon = \sigma^{-1}\sigma$. 这是明显的。

3）如果 σ 是可逆的，即存在线性变换 τ，使 $\sigma\tau = \varepsilon = \tau\sigma$， 那么$\sigma$必有逆变换。

所谓σ有逆变换，就是说σ是单满变换。

先说σ是单变换，假设 $\sigma(\alpha) = \sigma(\beta)$. 则有

$\tau(\sigma(\alpha)) = \tau(\sigma(\beta))$，$\tau\sigma(\alpha) = \tau\sigma(\beta)$，$\varepsilon(\alpha) = \varepsilon(\beta)$，$\alpha = \beta$.

再说σ是满变换，对任意的 $\alpha' \in V$. 令 $\tau(\alpha') = \alpha$，则有 $\sigma(\alpha) =$

$\sigma(\tau(\alpha')) = \sigma\tau(\alpha') = \varepsilon(\alpha') = \alpha'$.

4）如果σ是可逆的，那么使 $\sigma\tau = \varepsilon = \tau\sigma$ 的线性变换τ必定就是σ的逆变换，即 $\tau = \sigma^{-1}$.

因为σ是可逆的，由 3）σ 有逆变换 σ^{-1}. 于是 $\sigma\tau = \varepsilon = \sigma\sigma^{-1}$. 从而 $\sigma^{-1}(\sigma\tau) = \sigma^{-1}(\sigma\sigma^{-1})$，$(\sigma^{-1}\sigma)\tau = (\sigma^{-1}\sigma)\sigma^{-1}$，$\varepsilon\tau = \varepsilon\sigma^{-1}$，$\tau = \sigma^{-1}$.

总括起来，命题3给完全证明了，同时顺便说明：使$\sigma\tau = \varepsilon = \tau\sigma$的线性变换$\tau$是唯一的，即$\tau$是逆变换 σ^{-1}.

现在我们可以定义线性变换幂的概念及线性变换的多项式的概念。

设 $\sigma \in L(V)$. 由于线性变换的乘法满足结合律. 因而任意取定正整数 n，乘积

$$\overbrace{\sigma\ \sigma\ \cdots\ \sigma}^{n\text{个}}$$

都是一个确定的线性变换，叫做σ的n次幂，记作 $\sigma^n = \overbrace{\sigma\ \sigma\ \cdots\ \sigma}^{n\text{个}}$.

如果σ是可逆的，还可以定义负整数幂的概念，即规定

$$\sigma^{-n} = (\sigma^{-1})^n, \forall n \in N.$$

特别地，对任一线性变换 σ，我们总规定

$$\sigma^0 = \varepsilon.$$

根据线性变换幂的定义，可以推出指数法则：

$$\sigma^{m+n} = \sigma^m \cdot \sigma^n,\quad (\sigma^m)^n = \sigma^{mn},\quad m,\ n \geqslant 0;$$

当σ是可逆线性变换时，上述法则对任意的整数 m，n 都成立。但是，由于乘法不满足交换律，所以关于乘积的指数法则不成立，即

$$(\sigma\tau)^n \neq \sigma^n\tau^n.$$

设

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$$

是 $F[x]$ 中任一多项式，σ是V的任一线性变换，我们定义

$$f(\sigma) = a_m\sigma^m + a_{m-1}\sigma^{m-1} + \cdots + a_1\sigma + a_0\varepsilon$$

显然，$f(\sigma)$ 是一个线性变换，叫做线性变换σ的多项式。

不难验证，如果对 $F[x]$ 中多项式以下关系成立：

$$u(x)=f(x)+g(x),\quad v(x)=f(x)g(x)$$

那么对任一 $\sigma\in L(V)$，则有

$$u(\sigma)=f(\sigma)+g(\sigma),\quad v(\sigma)=f(\sigma)g(\sigma).$$

特别地

$$f(\sigma)g(\sigma)=g(\sigma)f(\sigma).$$

即同一个线性变换的多项式的乘法是可交换的。

例 2 在 $M_2(F)$ 中，转置变换

$$\tau:\ A\longmapsto A'$$

是单满变换，所以τ是可逆的，显然有 $\tau^2=\varepsilon$，从而 $\tau^{-1}=\tau$。再如线性变换

$$\sigma:\ \begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\longmapsto\begin{pmatrix}a_{11} & 0\\ 0 & a_{22}\end{pmatrix}$$

显然不是可逆的。

例 3 考虑平面$D^{(2)}$ 的两个线性变换

$$\sigma:\ (x,\ y)\longmapsto(-y,\ x),$$

$$\tau:\ (x,\ y)\longmapsto(0,\ y).$$

于是

$$\sigma\tau((x,\ y))=\sigma(\tau((x,\ y)))=\sigma((0,\ y))=(-y,\ 0),$$

$$\tau\sigma((x,\ y))=\tau(\sigma((x,\ y)))=\tau((-y,\ x))=(0,\ x).$$

显然，只要 $(x,\ y)\neq(0,\ 0)$，总有 $\sigma\tau((x,\ y))\neq\tau\sigma((x,\ y))$，这说明线性变换的乘法，一般是不满足交换律的。

这里还容易看出，τ 不是可逆的。但 σ 是一个单满变换，所以 σ 是可逆的，并且

$$\sigma^{-1}:\ (-y,x)\longmapsto(x,\ y),\ 即\ (x,\ y)\longmapsto(y,\ -x).$$

因为 $(y,\ -x)=-(-y,\ x)=-\sigma((x,\ y))$，从而 $\sigma^{-1}=-\sigma$，即 $\sigma^2=-\varepsilon$。

例 4 考虑无限维空间 $D[x]$，及其n维子空间 $D_n[x]$。我们

已经知道，求导数的方法 d 是空间 $D[x]$ 的一个线性变换．并且 $\forall f(x)\in D_n[x]$，都有

$$d(f(x))=f'(x)\in D_n[x]。即\ d(D_n[x])\subseteq D_n[x]。$$

这样，线性变换 d 在子空间 $D_n[x]$ 上有一个导出变换：$d|_{D_n[x]}$。我们把这个导出变换还记作 d．于是，$D_n[x]$ 中的多项式 $f(x)$，其次数 $<n$，所以 $f^{(n)}(x)=0$．这一事实说明，导出变换 d 具有性质：

$$d(f(x))=f'(x),\ d^2(f(x))=f''(x),\cdots,d^n(f(x))=f^{(n)}(x)=0.$$

即 d^n 是 $D_n[x]$ 的零变换：$d^n=0$。但是在整个空间 $D[x]$ 上看，对任意正整数 n，总有 $d^n\neq 0$。

例 5 在 $F_n[x]$ 中，关于 x 的平移变换

$$\sigma_a:\ f(x)\longmapsto f(x+a)$$

是一个线性变换。根据泰勒展开式

$$f(x+a)=f(x)+af'(x)+\frac{a^2}{2!}f''(x)+\cdots+\frac{a^{n-1}}{(n-1)!}f^{(n-1)}(x).$$

于是有

$$\begin{aligned}\sigma_a(f(x))&=f(x+a),\\ f(x)&=d^0(f(x)),\\ f'(x)&=d(f(x)),\\ f''(x)&=d^2(f(x)),\\ &\vdots\\ f^{(n-1)}(x)&=d^{n-1}(f(x)),\end{aligned}$$

从而就有

$$\begin{aligned}\sigma_a(f(x))&=d^0(f(x))+ad(f(x))+\frac{a^2}{2!}d^2(f(x))+\cdots\\ &\quad+\frac{a^{n-1}}{(n-1)!}d^{n-1}(f(x))\\ &=\left(d^0+ad+\frac{a^2}{2!}d^2+\cdots+\frac{a^{n-1}}{(n-1)!}d^{n-1}\right)(f(x)),\end{aligned}$$

这样，按照变换相等的定义，即得

$$\sigma_a = \varepsilon + ad + \frac{a^2}{2!}d^2 + \cdots + \frac{a^{n-1}}{(n-1)!}d^{n-1},$$

这个等式说明，线性变换σ_a是线性变换d的一个多项式。

练　习　二

1．在几何空间$D^{(3)}$中，取直角坐标系$OXYZ$，以σ表示将空间绕OX轴由OY向OZ方向旋转90°的变换，以τ表示绕OY轴由OZ向OX方向旋转90°的变换，以ρ表示绕OZ轴由OX向OY方向旋转90°的变换，证明

1）$\sigma^4=\tau^4=\rho^4=\varepsilon$

2）$\sigma\tau\neq\tau\sigma$

3）$\sigma^2\tau^2=\tau^2\sigma^2$

4）$(\sigma\tau)^2\neq\sigma^2\tau^2$

2．在$M_2(F)$中，设线性变换

$$\sigma\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$$

$$\tau\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}\lambda a & 0\\ 0 & \mu d\end{pmatrix}$$

求 $\sigma+\tau$，$\sigma\tau$。

3．在数域F上的一元多项式构成的线性空间$F[x]$中，如果

$$\sigma(f(x))=f'(x)$$

$$\tau(f(x))=xf(x)$$

证明：

$$\sigma\tau-\tau\sigma=\varepsilon \qquad (\varepsilon\text{ 是恒等变换})$$

4．设 σ，τ是线性变换，如果

$$\sigma\tau-\tau\sigma=\varepsilon$$

证明：

$$\sigma^k\tau-\tau\sigma^k=k\sigma^{k-1}\qquad (k>1)$$

5．设 σ，τ 是线性变换，且

$$\sigma^2=\sigma,\quad \tau^2=\tau$$

证明：

1）若 $(\sigma+\tau)^2=\sigma+\tau$，则 $\sigma\tau=\theta$ （θ是零变换）

2）若 $\sigma\tau=\tau\sigma$，则 $(\sigma+\tau-\sigma\tau)^2=\sigma+\tau-\sigma\tau$。

6．设 σ，τ是V的线性变换，定义

$$[\sigma,\ \tau]=\sigma\tau-\tau\sigma$$

证明：对任意 σ，τ，ρ 以下等式成立。

$$[[\rho,\ \sigma],\ \tau]+[[\sigma,\ \tau],\rho]+[[\tau,\ \rho],\ \sigma]=\theta$$

7．可逆线性变换的乘积仍是可逆线性变换。

8．设σ是线性空间V的可逆线性变换，α_1，α_2，…，α_n 是线性无关的向量组，证明

$$\sigma(\alpha_1),\ \sigma(\alpha_2),\ \cdots,\ \sigma(\alpha_n)$$

也是线性无关的。

§3　线性变换与子空间的关系

设σ是线性空间V的线性变换，W是V的子空间。于是与σ，W相联系的有以下两个子集：

$$\sigma(W)=\{\sigma(\xi)\ |\ \xi\in W\};\quad \sigma^{-1}(W)=\{\xi\in V\ |\ \sigma(\xi)\in W\}$$

称$\sigma(W)$为W在σ之下的象，$\sigma^{-1}(W)$ 为W在σ之下的原象。

命题1　设σ是V的线性变换，W是V的子空间，那么 $\sigma(W)$与$\sigma^{-1}(W)$ 都是V的子空间。

证　任取 α'，$\beta'\in\sigma(W)$，于是必有 α，$\beta\in W$，使$\sigma(\alpha)=\alpha'$，$\sigma(\beta)=\beta'$。由此即得

$$k\alpha'+h\beta'=k\sigma(\alpha)+h\sigma(\beta)=\sigma(k\alpha+h\beta),$$

因为W是子空间，可知 $k\alpha+h\beta\in W$。从而便有

$$k\alpha'+h\beta'\in\sigma(W)$$

所以 $\sigma(W)$ 是子空间。

再有，对任意的 α，$\beta\in\sigma^{-1}(W)$，即 $\sigma(\alpha)$，$\sigma(\beta)\in W$。于是 $\sigma(k\alpha+h\beta)=k\sigma(\alpha)+h\sigma(\beta)\in W$，这就表明了，$k\alpha+h\beta\in\sigma^{-1}(W)$，即 $\sigma^{-1}(W)$ 是子空间。

特别地，把子空间 $\sigma(V)$ 叫做σ的值域；把子空间 $\sigma^{-1}(0)$ 叫做σ的核。

不难指出，σ 是满变换当且仅当 $\sigma(V)=V$；σ 是单变换当且仅当 $\sigma^{-1}(0)=\{0\}$。前者是自明的，下面对后者解释几句，如果 σ 是单的，自然有 $\sigma^{-1}(0)=\{0\}$。如果σ不是单的，那么只少有两个不同的向量 α，β使 $\sigma(\alpha)=\sigma(\beta)$。由此便有

$$\alpha-\beta\neq 0,\ \text{但}\ \sigma(\alpha-\beta)=\sigma(\alpha)-\sigma(\beta)=0.$$

这就说明核 $\sigma^{-1}(0)\neq\{0\}$。

例 1　考虑线性空间 $F_n[x]$ 及求导数的线性变换 d。于是

d 的值域 $d(F_n[x])=F_{n-1}[x]$，d 的核 $d^{-1}(0)=F_1[x]$。值得注意的是，这个求导数的线性变换 d，如果在线性空间 $F[x]$ 里来考虑它的作用的话，d 的核还是 $F_1[x]$，但d的值域是全空间：$d(F[x])=F[x]$。这就是说：d 是 $F[x]$ 的满变换，但不是$F_n[x]$的满变换。

例 2　在三维几何空间$D^{(3)}$ 中，考虑一个射影变换

$$\sigma:\ (x,\ y,\ z)\longmapsto(x,\ y,0)$$

于是σ的值域就是 xy——平面：$\sigma(D^{(3)})=\{(x,\ y,0)\mid x,\ y\in D\}$。$\sigma$的核就是$z$——轴：$\sigma^{-1}(0)=\{0,\ 0,\ z)\mid z\in D\}$。

命题 2　若子空间 $W=L(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_s)$，那么$W$在$\sigma$之下的象 $\sigma(W)=L(\sigma(\alpha_1),\ \sigma(\alpha_2),\ \cdots,\ \sigma(\alpha_s))$。换句话说就是，生成组的象是象空间的生成组。

证明　首先自然有 $\sigma(\alpha_1)$，$\sigma(\alpha_2)$，$\cdots$，$\sigma(\alpha_s)\in\sigma(W)$。其次，任取 $\sigma(W)$ 的一个向量 α'。于是有 $\alpha\in W$，使 $\sigma(\alpha)=\alpha'$。把向量 α 用W的生成组 α_1，α_2，$\cdots$，α_s 表示出来，即有

$$\alpha=k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s.$$

这样便得

$$\alpha' = \sigma(\alpha) = \sigma(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s) = k_1\sigma(\alpha_1) + k_2\sigma(\alpha_2) + \cdots + k_s\sigma(\alpha_s).$$

即 $\sigma(W) = L(\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_s))$.

推论 $\dim\sigma(W) \leqslant \dim W$.

事实上 $\{\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_s)\}$ 的秩$\leqslant\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ 的秩，即 $\dim\sigma(W) \leqslant \dim W$.

命题 3 **设W_1，W_2 是V 的子空间，σ 是V 的线性变换．那么 $\sigma(W_1 + W_2) = \sigma(W_1) + \sigma(W_2)$.**

事实上对任意的 $\alpha_1 \in W_1$，$\alpha_2 \in W_2$，等式

$$\sigma(\alpha_1 + \alpha_2) = \sigma(\alpha_1) + \sigma(\alpha_2)$$

就是命题 3 所要的结论．

命题 4 **设W_1，W_2 是V 的子空间，σ 是V 的线性变换．那么 $\sigma(W_1 \cap W_2) \subseteq \sigma(W_1) \cap \sigma(W_2)$.**

事实上 由 $W_1 \cap W_2 \subseteq W_1$，$W_1 \cap W_2 \subseteq W_2$ 便得

$$\sigma(W_1 \cap W_2) \subseteq \sigma(W_1), \quad \sigma(W_1 \cap W_2) \subseteq \sigma(W_2)$$

所以 $\sigma(W_1 \cap W_2) \subseteq \sigma(W_1) \cap \sigma(W_2)$.

推论 **如果σ是可逆的，则有**

$$\sigma(W_1 \cap W_2) = \sigma(W_1) \cap \sigma(W_2).$$

事实上 $\sigma(W_1 \cap W_2) \subseteq \sigma(W_1) \cap \sigma(W_2)$．为了推出

$$\sigma(W_1) \cap \sigma(W_2) \subseteq \sigma(W_1 \cap W_2)$$

我们考虑σ的逆变换 σ^{-1}．由命题 4，则有

$$\sigma^{-1}(\sigma(W_1) \cap \sigma(W_2)) \subseteq \sigma^{-1}(\sigma(W_1)) \cap \sigma^{-1}(\sigma(W_2))$$

显然有 $\sigma^{-1}(\sigma(W_1)) = W_1$，$\sigma^{-1}(\sigma(W_2)) = W_2$，于是

$$\sigma^{-1}(\sigma(W_1) \cap \sigma(W_2)) \subseteq W_1 \cap W_2$$

从而可得

$$\sigma(W_1) \cap \sigma(W_2) \subseteq \sigma(W_1 \cap W_2).$$

这就证明了：$\sigma(W_1 \cap W_2) = \sigma(W_1) \cap \sigma(W_2)$.

我们明确指出，如果 σ 不是可逆的，上述推论是不成立的．例

如，$D^{(2)}$ 的子空间

$$W_1=\{(x, x) \mid x\in D\}, W_2=\{(x, 0) \mid x\in D\}.$$

线性变换

$$\sigma: (x, y) \longmapsto (x, 0).$$

于是

$W_1\cap W_2=\{0\}$, $\sigma(W_1)=W_2$, $\sigma(W_2)=W_2$, $\sigma(W_1)\cap\sigma(W_2)=W_2$。所以 $\sigma(W_1\cap W_2)=\{0\}\neq\sigma(W_1)\cap\sigma(W_2)$。

这里再顺便说明一个问题，容易看出：

$$D^{(2)}=W_1\dot{+}W_2$$

但 $\sigma(W_1\dot{+}W_2)=\sigma(W_1)+\sigma(W_2)$，而 $\sigma(W_1)\cap\sigma(W_2)=W_2$，所以 $\sigma(W_1)+\sigma(W_2)$ 不是直和。这说明直和的象不一定还是直和。

练　习　三

1. 设σ是 $V_n(F)$ 中的线性变换，如果

$$\alpha_1, \alpha_2, \cdots, \alpha_s, \alpha_{s+1}, \cdots, \alpha_n$$

是$V_n(F)$ 的基底，且 α_1, α_2, $\cdots$, α_s 是 $\sigma^{-1}(0)$ 的基底，证明：

1) $\sigma(\alpha_{s+1})$, $\sigma(\alpha_{s+2})$, $\cdots$, $\sigma(\alpha_n)$ 是 $\sigma(V_n(F))$ 的基底。

2) $\dim\sigma^{-1}(0)+\dim\sigma(V_n(F))=n$

2. 由上题知线性变换 σ 的核空间的维数与象空间的维数和等于线性空间的维数。据此是否可以断言

$$\sigma^{-1}(0)+\sigma(V_n(F))=V_n(F)$$

为什么？

3. 设σ是n维线性空间V 的线性变换，证明σ 是可逆线性变换的充分必要条件是 $\sigma^{-1}(0)=\{0\}$

4. 设σ 是 $V_n(F)$ 的一个线性变换，W是$V_n(F)$ 的子空间，证明：

$$\dim(\sigma(W))+\dim(\sigma^{-1}(0)\cap W)=\dim(W)$$

5. 设σ, τ是 $V_n(F)$ 的线性变换，证明

$$\dim(\sigma\tau(V)) \geqslant \dim(\sigma(V)) + \dim(\tau(V)) - n$$

6. 设 σ，τ是$V_n(F)$ 的线性变换，证明

$$\dim(\sigma\tau)^{-1}(0)) \leqslant \dim(\sigma^{-1}(0)) + \dim(\tau^{-1}(0))$$

§4 不变子空间

在线性变换的理论中，不变子空间是个基本的概念。本节介绍不变子空间的简单性质以及它对化简线性变换的作用。

定义 1 设σ是V的线性变换，W是V的子空间，如果$\sigma(W) \subseteq W$，则称W是σ的不变子空间。

例 1 在三维几何空间$D^{(3)}$ 中，线性变换

$$\sigma: (x, y, z) \longmapsto (x, y, 0),$$

子空间$W_1 = \{(0, y, z) \mid y, z \in D\}$，$W_2 = \{(0, y, y) \mid y \in D\}$。于是

$$\sigma(W_1) \subseteq W_1, \quad \sigma(W_2) \not\subseteq W_2$$

所以，W_1 是σ的不变子空间，W_2不是σ的不变子空间。

例 2 在$D[x]$ 中，求导数的变换

$$d: f(x) \longmapsto f'(x),$$

子空间$D_n[x]$ 是d的不变子空间。但是对于求积分的变换

$$s: f(x) \longmapsto \int_0^x f(t)\,dt$$

来说，$D_n[x]$ 则不是s的不变子空间。

例 3 在$M_n(F)$ 中，转置变换

$$\tau: A \longmapsto A'.$$

子空间$W_1 = \{A \mid A' = A\}$，$W_2 = \{A \mid A' = -A\}$。于是W_1与W_2都是τ的不变子空间。

例 4 V是任一线性空间，σ 是V的任一线性变换，于是零空间总是σ的不变子空间。V本身做为V的子空间，也总是σ的不变子空间。称这两个特殊的不变子空间为当然的不变子空间或平凡的不

变子空间。

例 5　考虑二维几何空间$D^{(2)}$，线性变换

$$\sigma:\ (x,\ y)\longmapsto\left(\frac{1}{\sqrt{2}}x+\frac{-1}{\sqrt{2}}y,\ \ \frac{1}{\sqrt{2}}x+\frac{1}{\sqrt{2}}y\right).$$

这个σ就是逆时针旋转$\frac{\pi}{4}$的旋转变换。容易想到，σ 除了当然的不变子空间之外，再没有不变子空间。

例 6　设V是任一线性空间，线性变换

$\sigma:\ \alpha\longmapsto k\alpha,\quad k$为$F$中确定的数。

显然，对V的任一子空间W来说都有 $\sigma(W)\subseteq W$，即W是σ 的不变子空间。

命题 1　若 W_1，W_2 都是σ的不变子空间，那么 $W_1\cap W_2$ 与 W_1+W_2 也都是σ的不变子空间。

证明　任取$\alpha\in W_1\cap W_2$，即 $\alpha\in W_1$，$\alpha\in W_2$。从而$\sigma(\alpha)\in W_1$，$\sigma(\alpha)\in W_2$，即有 $\sigma(\alpha)\in W_1\cap W_2$。所以 $W_1\cap W_2$ 是σ 的不变子空间。同理可证 W_1+W_2 也是σ的不变子空间。

下面探讨不变子空间在化简线性变换方面的意义。

设W是σ的不变子空间。这时σ在W上导出一个线性变换 $\sigma|_W$。有时也称σ在W上的导出变换为σ在W上的限制，用以体现仅仅考虑σ 在W上的作用。

例 7　在三维几何空间$D^{(3)}$ 中，考虑线性变换

$\tau:\ (x,\ y,\ z)\longmapsto(0,\ y,\ z)$。

于是例 1 里提到的两个子空间

$W_1=\{(0,\ y,\ z)\ |\ y,\ z\in D\}$，　$W_2=\{(0,\ y,\ y)\ |\ y\in D\}$

都是τ的不变子空间。这时τ在W_1与W_2 上各自导出一个线性变换，τ 在W_1 上的导出变换$\tau|_{W_1}$ 是 W_1 的恒等变换，τ 在W_2 上的导出变换 $\tau|_{W_2}$ 也是恒等变换。再如子空间 $W_3=\{(x,\ 0,\ 0)\ |\ x\in D\}$ 也是τ的不变子空间，τ在W_3 上的限制$\tau|_{W_3}$ 是零变换。

例 8　$M_n(F)$的转置变换τ，如例 3 中的两个不变子空间W_1 与

W_2。于是 $\tau|_{W_1}=\varepsilon$，$\tau|_{W_2}=-\varepsilon$。即 $\tau|_{W_1}$ 是 W_1 的恒等变换 ε，$\tau|_{W_2}$ 是 W_2 的恒等变换的 -1 倍 $-\varepsilon$。

例 9　考虑 $F[x]$ 的求导数的变换 d，子空间 $F_n[x]$。于是 $d|_{F_n[x]}=d$，即 d 在 $F_n[x]$ 上导出的变换就是（$F_n[x]$ 的）求导数的变换 d。

进而考虑，如果线性空间 V 能够分解成两个子空间的直和：

$$V=W_1\dotplus W_2,$$

并且 W_1，W_2 都是线性变换 σ 的不变子空间。于是 σ 在这两个不变子空间上各有一个导出变换 $\sigma|_{W_1}$，$\sigma|_{W_2}$。这样，对 V 中任一向量 α，有唯一的分解式

$$\alpha=\alpha_1+\alpha_2,\quad \alpha_1\in W_1,\quad \alpha_2\in W_2.$$

则有

$$\begin{aligned}\sigma(\alpha)&=\sigma(\alpha_1)+\sigma(\alpha_2)\\&=\sigma|_{W_1}(\alpha_1)+\sigma|_{W_2}(\alpha_2)\end{aligned}$$

这时，如果我们记作 $\sigma=\sigma|_{W_1}+\sigma|_{W_2}$，即有

$$\sigma(\alpha)=(\sigma|_{W_1}+\sigma|_{W_2})(\alpha)=\sigma|_{W_1}(\alpha_1)+\sigma|_{W_2}(\alpha_2).$$

以上论述表明：当线性空间 V 分解成关于线性变换 σ 的两个不变子空间的直和：$V=W_1\oplus W_2$ 时，线性变换 σ 对全空间 V 的作用便可分解为两个导出变换 $\sigma|_{W_1}$，$\sigma|_{W_2}$ 分别在 W_1 与 W_2 上作用的和。

反之，如果 $V=W_1\oplus W_2$。σ_1，σ_2 分别是线性空间 W_1 与 W_2 的线性变换。那么我们可以得到 V 的一个确定的线性变换 σ，以 W_1，W_2 为不变子空间，并且 σ 在 W_1 上的导出变换是 σ_1：$\sigma|_{W_1}=\sigma_1$；在 W_2 上的导出变换是 σ_2：$\sigma|_{W_2}=\sigma_2$。

事实上对 V 中任一向量 $\alpha=\alpha_1+\alpha_2$，$\alpha_1\in W_1$，$\alpha_2\in W_2$。我们自然的规定 V 的一个变换

$$\sigma(\alpha)=\sigma_1(\alpha_1)+\sigma_2(\alpha_2).$$

不难验证 σ 是线性的，W_1 与 W_2 都是 σ 的不变子空间。而且明显的，$\sigma|_{W_1}=\sigma_1$，$\sigma|_{W_2}=\sigma_2$。

综合以上两个方面的分析，即得：如果线性空间 V 能够分解成

关于σ不变的两个子空间W_1与W_2的直和，那么对于V的线性变换σ的研究归结为对线性空间W_1与W_2的线性变换的研究。一般地，不难想到：如果$V = W_1 \dotplus W_2 \dotplus \cdots \dotplus W_s$，$W_i$都是$\sigma$的不变子空间，那么对$V$的线性变换$\sigma$的研究归结为对各个子空间$W_1, W_2, \cdots, W_s$的线性变换的研究。

例10 在$M_2(F)$中，τ为转置变换，W_1与W_2为例3中提到的子空间，它们都是τ的不变子空间，而且使

$$M_2(F) = W_1 \dotplus W_2.$$

于是转置变换τ分解为导出变换的和：

$$\tau = \tau|_{W_1} + \tau|_{W_2}$$

这里 $\tau|_{W_1} = \varepsilon$，$\tau|_{W_2} = -\varepsilon$。所以关于二阶方阵转置变换的研究归结为对称阵上的恒等变换与反对称阵上的，以-1为系数的，相似变换的研究。

本节最后我们给出与一维不变子空间密切联系着的重要概念——线性变换的特征向量与特征值。

设W是σ的一维不变子空间。于是必有向量$\xi \in W$，使$W = [\xi]$。因为W是σ的不变子空间，必有

$$\sigma(\xi) = \lambda_0 \xi, \quad \xi \neq 0. \tag{1}$$

反之，V中有非零向量ξ使（1）式成立，那么一维子空间$W = [\xi]$是σ的不变子空间。这样，σ有一维不变子空间当且仅当有非零向量ξ，使（1）式成立。

定义2 设σ是V的线性变换，$\xi \in V$，$\xi \neq 0$。如果

$$\sigma(\xi) = \lambda_0 \xi,$$

则称向量ξ是σ的属于特征值λ_0的特征向量，而λ_0叫做σ的特征值。

例如，$M_2(F)$的转置变换τ，则有

$$\tau\left(\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}\right) = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \quad \tau\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) = -\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

这样，$M_2(F)$中的向量$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$是$\tau$的属于特征值1的特征向量；

$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$是属于特征值$-1$的特征向量.

命题2　设 α, β 都是σ的属于同一个特征值 λ_0 的特征向量，那么线性组合 $k\alpha+h\beta$ 不是零向量时就还是σ的属于特征值λ_0的特征向量.

证明　假设 $k\alpha+h\beta\neq 0$. 于是

$$\begin{aligned}\sigma(k\alpha+h\beta)&=k\sigma(\alpha)+h\sigma(\beta)\\&=k(\lambda_0\alpha)+h(\lambda_0\beta)\\&=\lambda_0(k\alpha+h\beta).\end{aligned}$$

即 $k\alpha+h\beta$ 也是σ的属于特征值λ_0的特征向量.

特别地，若 $\xi\neq 0$, $\sigma(\xi)=\lambda_0\xi$, 那么 $W=[\xi]$ 中每一个非零向量都是σ的属于特征值λ_0的特征向量.

推论　σ 的属于同一个特征值λ_0的全部特征向量再添上零向量组成的集合 $S_\sigma(\lambda_0)$ 是σ的一个不变子空间. 其实就是

$$S_\sigma(\lambda_0)=\{\xi\in V\mid\sigma(\xi)=\lambda_0\xi\}.$$

我们称$S_\sigma(\lambda_0)$为σ的属于特征值λ_0的特征子空间.

例11　考虑三维几何空间$D^{(3)}$ 的线性变换

$$\sigma:\ (x,\ y,\ z)\longrightarrow(x,\ y,\ 0).$$

于是 $W_1=[(0,\ 0,\ 1)]$ 是 σ 的一个一维不变子空间，其中每一非零向量都是 σ 的属于特征值 0 的特征向量. 反之，任一属于特征值 0 的特征向量也都被包含在W_1里，所以W_1恰是σ的属于特征值 0 的特征子空间 $S_\sigma(0)$. 此外，1 也是σ的特征值. 向量$\xi_1=(1, 0, 0)$就是一个属于特征值 1 的特征向量：$\sigma(\xi_1)=1\cdot\xi_1$. 这样 $W_2=[\xi_1]$ 也是 σ 的一个一维不变子空间，当然，其中的每一个非零向量都是 σ 的属于特征值 1 的特征向量. 我们容易发现，除了 W_2 中的向量，$D^{(3)}$ 里还有σ的属于特征值 1 的特征向量. 如 $\xi_2=(0,\ 1,\ 0)$ 就是属于特征值 1 的特征向量：$\sigma(\xi_2)=1\cdot\xi_2$. 这就是说，$W_2$ 中还没全包含属于特征值 1 的一切特征向量，所以，W_2不是σ的属于特征值 1 的特征子空间. 可是，$W_3=[\xi_1,\ \xi_2]$ 已经包含一切属于特征值

1的特征向量，即 $W_3=S_\sigma(1)$。

命题3　设 λ_1，λ_2 是σ的两个不同的特征值。ξ_1，ξ_2 分别是属于 λ_1，λ_2 的特征向量，那么 ξ_1，ξ_2 是线性无关的。

证明　考察线性表出式

$$k_1\xi_1+k_2\xi_2=0. \qquad (2)$$

用σ作用上式两端的向量，则有

$$\sigma(k_1\xi_1+k_2\xi_2)=0,$$

即

$$k_1\sigma(\xi_1)+k_2\sigma(\xi_2)=0$$

$$k_1\lambda_1\xi_1+k_2\lambda_2\xi_2=0. \qquad (3)$$

用λ_1乘（2）式减去（3）式，则得

$$(\lambda_1-\lambda_2)k_2\xi_2=0$$

因为 $\xi_2\neq0$，则有 $(\lambda_1-\lambda_2)k_2=0$。又因 $\lambda_1\neq\lambda_2$，即 $\lambda_1-\lambda_2\neq0$，便得 $k_2=0$。将 $k_2=0$ 代入（2）式又得 $k_1=0$。这就证明了 ξ_1，ξ_2 是线性无关的。

练　习　四

1．若W是σ，τ 的不变子空间，则W也是 $\sigma+\tau$，$\sigma\tau$ 的不变子空间。

2．设σ是n维线性空间V的可逆线性变换，如果W是σ的不变子空间，那么W也是σ^{-1} 的不变子空间。

3．设 $V_n(F)$ 中的线性变换 σ，τ 是可换的，证明 $\tau^{-1}(0)$，$\tau(V)$ 是σ的不变子空间。

4．设W是σ的不变子空间，则 $\sigma(W)$ 也是σ的不变子空间。W在σ之下的完全原象 $\sigma^{-1}(W)$ 也是σ的不变子空间。

5．证明$V_n(F)$ 的线性变换 σ 是相似变换的充分必要条件是$V_n(F)$ 中任一一维子空间都是σ的不变子空间。

6．设 $W=L(\alpha_1, \alpha_2, \cdots, \alpha_m)$，证明$W$是线性变换 σ 的不变

子空间的充分必要条件是

$$\sigma(\alpha_1),\ \sigma(\alpha_2),\ \cdots,\ \sigma(\alpha_m) \in W.$$

§5 有限维空间的线性变换

对于有限维空间来说，每一个线性变换都可以用一组数具体的表示出来。这样就使n维空间上的线性变换与n阶方阵联系起来，进而使得线性方程组的理论及n阶方阵相似标准形理论成为讨论有限维空间线性变换问题的有力的代数方法。

设V是F上n维线性空间，σ是V的线性变换。任意取定V的一个基底 ε_1，ε_2，$\cdots$，ε_n。对任一向量

$$\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n$$

则有

$$\sigma(\xi) = x_1\sigma(\varepsilon_1) + x_2\sigma(\varepsilon_2) + \cdots + x_n\sigma(\varepsilon_n) \qquad (1)$$

上式表明，任一向量ξ在σ之下的象都可表成基底向量 ε_1，ε_2，$\cdots$，ε_n 在σ之下的象的线性组合。进而把 $\sigma(\varepsilon_1)$，$\sigma(\varepsilon_2)$，$\cdots$，$\sigma(\varepsilon_n)$ 用基底 ε_1，ε_2，$\cdots$，ε_n 线性表示出来，即有

$$\begin{aligned}
\sigma(\varepsilon_1) &= a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots + a_{n1}\varepsilon_n \\
\sigma(\varepsilon_2) &= a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots + a_{n2}\varepsilon_n \\
&\cdots\quad \cdots\quad \cdots \\
\sigma(\varepsilon_n) &= a_{1n}\varepsilon_1 + a_{2n}\varepsilon_2 + \cdots + a_{nn}\varepsilon_n
\end{aligned}$$

令

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$

于是

$$(\sigma(\varepsilon_1)\sigma(\varepsilon_2)\ \cdots\sigma(\varepsilon_n)) = (\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)A$$

如果简记 $(\sigma(\varepsilon_1)\sigma(\varepsilon_2)\cdots\sigma(\varepsilon_n)) = \sigma(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)$

则有

$$\sigma(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n) = (\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)A$$

令

$$\sigma(\xi) = \eta = y_1\varepsilon_1 + y_2\varepsilon_2 + \cdots + y_n\varepsilon_n$$

即ξ在σ之下的象η在基底 ε_1, ε_2, …, ε_n 上的坐标为 y_1, y_2, …, y_n。下面我们指出ξ的坐标与象η的坐标之间的关系。

$$\sigma(\xi) = \sigma(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

$$= (\varepsilon_1\ \varepsilon_2\cdots\ \varepsilon_n)A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

从而

$$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix} = A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix} \qquad (2)$$

称（2）式为线性变换σ在基底 ε_1, ε_2, …, ε_n 上的坐标表示式。其中的n阶方阵A叫做σ在基底 ε_1, ε_2, …, ε_n 上的**表示矩阵**。

综上所述即有：任意取定n维空间V的一个基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 之后，V的每一线性变换σ都有唯一的表示矩阵A，使

$$\sigma(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n) = (\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)A,$$

并且，如果

$$\xi = (\varepsilon_1\ \varepsilon_2\cdots\varepsilon_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad \sigma(\xi) = \eta = (\varepsilon_1\ \varepsilon_2\cdots\varepsilon_n)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$$

那么

$$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix} = A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$

定理 设V是n维线性空间，任取一个基底 ε_1, ε_2, …, ε_n。那

么V的任一线性变换都有唯一的表示矩阵．反之，任意给定一个n阶方阵A，都有唯一的线性变换σ，在基底ε_1，ε_2，…，ε_n上的表示矩阵就是A．

证明　定理的头一个结论已如上述．定理的第二个结论也是明显的．这只要利用公式

$$A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}，\text{其中}\ \xi=(\varepsilon_1\ \varepsilon_2\cdots\ \varepsilon_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

$$\eta=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}，\text{规定}V\text{的一个线性变换}$$

$$\sigma(\xi)=\eta,$$

那么σ在基底ε_1，ε_2，…，ε_n上的表示矩阵就是A，而在基底ε_1，ε_2，…，ε_n上以A为表示矩阵的线性变换当然只有一个．

这个定理表明，通过取定一个基底，使n维空间的线性变换的集合与n阶方阵的集合之间建立起一个一一对应关系．

命题1　**设σ，τ是$V=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$的两个线性变换．在基底$\varepsilon_1$，$\varepsilon_2$，…，$\varepsilon_n$上的表示矩阵分别为$A$与$B$．那么**

1）$\sigma+\tau$的表示矩阵为$A+B$，

2）$k\sigma$的表示矩阵为kA，

3）$\sigma\tau$的表示矩阵为AB，

4）σ是可逆的当且仅当A是可逆的，并且σ^{-1}的表示矩阵为A^{-1}．

证明　由σ，τ的表示矩阵分别为A，B．即有

$$\sigma(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)=(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)A$$

$$\tau(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)=(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)B.$$

于是

$$(\sigma+\tau)(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)=\sigma(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)\ +\tau(\varepsilon_1\ \varepsilon_2\ \cdots\varepsilon_n)$$

$$
\begin{aligned}
&= (\varepsilon_1\ \varepsilon_2 \cdots \varepsilon_n)A + (\varepsilon_1\ \varepsilon_2\ \cdots \varepsilon_n)B \\
&= (\varepsilon_1\ \varepsilon_2 \cdots \varepsilon_n)(A+B),
\end{aligned}
$$

$$
\begin{aligned}
(k\sigma)(\varepsilon_1\ \varepsilon_2 \cdots \varepsilon_n) &= k\sigma(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n) \\
&= k(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)A \\
&= (\varepsilon_1\ \varepsilon_2 \cdots \varepsilon_n)kA,
\end{aligned}
$$

$$
\begin{aligned}
\sigma\tau(\varepsilon_1\ \varepsilon_2 \cdots \varepsilon_n) &= \sigma[\tau(\varepsilon_1\ \varepsilon_2 \cdots \varepsilon_n)] \\
&= \sigma((\varepsilon_1\ \varepsilon_2\ \cdots \varepsilon_n)B) \\
&= [\sigma(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)]B \\
&= (\varepsilon_1\ \varepsilon_2\ \cdots \varepsilon_n)AB
\end{aligned}
$$

最后，如果σ是可逆的，令σ^{-1}的表示矩阵为C.于是由$\sigma\sigma^{-1}=\varepsilon$可得$AC=I$，从而$A$是可逆的，并且$\sigma^{-1}$的表示矩阵就是$A^{-1}$。反之，如果$\sigma$的表示矩阵$A$是可逆的，令$\rho$是以$A^{-1}$为表示矩阵的线性变换。于是由$AA^{-1}=I$可得$\sigma\rho=\varepsilon=\rho\sigma$，从而$\sigma$是可逆的。

命题2　设$W=L(\alpha_1, \alpha_2, \cdots, \alpha_s)$，$\sigma$的表示矩阵为$A$。那么

$$\dim\sigma(W)=\operatorname{rank}A(\bar{\alpha}_1\ \bar{\alpha}_2 \cdots \bar{\alpha}_s)$$

其中$(\bar{\alpha}_1\ \bar{\alpha}_2 \cdots \bar{\alpha}_s)$是$\alpha_1, \alpha_2, \cdots, \alpha_s$的坐标列矩阵。

证明

$$
\begin{aligned}
\dim\sigma(W) &= \operatorname{rank}\{\sigma(\alpha_1), \sigma(\alpha_2), \cdots, (\alpha_s)\} \\
&= \operatorname{rank}(A\bar{\alpha}_1\ A\bar{\alpha}_2 \cdots A\bar{\alpha}_s) \\
&= \operatorname{rank}A(\bar{\alpha}_1\ \bar{\alpha}_2 \cdots \bar{\alpha}_s).
\end{aligned}
$$

命题3　$\dim\sigma(V)=\operatorname{rank}A$, $\dim\sigma^{-1}(0)=n-\operatorname{rank}A$.

证明　令$V=[\alpha_1, \alpha_2, \cdots, \alpha_n]$，这时$(\bar{\alpha}_1\ \bar{\alpha}_2 \cdots \bar{\alpha}_n)$是单位矩阵，于是由命题2即得$\dim\sigma(V)=\operatorname{rank}A$.

再有$\sigma^{-1}(0)=\{\xi\in V \mid \sigma(\xi)=0\}$。于是

$$\xi\in\sigma^{-1}(0) \text{ 当且仅当 } A\bar{\xi}=0.$$

这表明核$\sigma^{-1}(0)$里的向量ξ的坐标列$\bar{\xi}$是齐次线性方程组

$$A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

的解，反之亦然．换句话说就是，$\sigma^{-1}(0)$ 中向量的坐标列恰是上述齐次方程组的解空间．所以，$\dim\sigma^{-1}(0)=n-\text{rank}A$．

推论　若σ是满变换，则σ就是可逆的；若σ是单变换，则σ就是可逆的．

事实上　若σ是满的，则有$\dim\sigma(V)=\text{rank}A=n$．从而$\dim\sigma^{-1}(0)=n-\text{rank}A=0$，所以$\sigma$又是单的，即$\sigma$是可逆的．若$\sigma$是单的，则有 $\dim\sigma^{-1}(0)=n-\text{rank}A=0$．从而 $\dim\sigma(V)=\text{rank}A=n$．所以$\sigma$又是满的，即$\sigma$是可逆的．

我们称 $\sigma(V)$ 的维数为σ的**秩**，$\sigma^{-1}(0)$ 的维数为σ 的**亏数**．显然，线性变换σ的秩等于它的表示矩阵A的秩．

命题 4　**设 $V=[\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n]$，$\sigma$ 在 $\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n$ 上的表示矩阵为A．那么**

$\sigma\xi=\lambda_0\xi$ **当且仅当** $(\lambda_0\varepsilon-\sigma)\xi=0$ **当且仅当** $(\lambda_0 I-A)\overline{\xi}=0$．

命题 4 是很明显的，证明从略．

这个命题 4 提供一个求线性变换 σ 的特征值及相应的特征向量的方法．即σ的特征值就是其表示矩阵A的特征根，而σ的特征向量的坐标列恰是表示矩阵A的特征向量．

例 1　$D^{(3)}=[\varepsilon_1,\ \varepsilon_2,\ \varepsilon_3]$，$\varepsilon_1=(1,\ 0,\ 0)$，$\varepsilon_2=(0,\ 1,\ 0)$，$\varepsilon_3=(0,\ 0,\ 1)$．考察下列变换：

$$\sigma(\xi)=\sigma((x,\ y,\ z))=(x,\ y,\ 0),$$
$$\tau(\xi)=\tau((x,\ y,\ z))=(-x,\ y,\ z),$$
$$\rho(\xi)=\rho((x,\ y,\ z))=(x+y,\ y+z,\ z+x).$$

容易知道 $\sigma,\ \tau,\ \rho$ 都是$D^{(3)}$ 的线性变换，求出它们在 $\varepsilon_1,\ \varepsilon_2,\ \varepsilon_3$ 上的表示矩阵，进而再求 $\sigma+\tau,\ \tau\rho$ 的表示矩阵．

解　由

$$\sigma(\varepsilon_1)=(1,\ 0,\ 0)\ =1\cdot\varepsilon_1+0\cdot\varepsilon_2+0\cdot\varepsilon_3$$
$$\sigma(\varepsilon_2)=(0,\ 1,\ 0)=0\cdot\varepsilon_1+1\cdot\varepsilon_2+0\cdot\varepsilon_3$$
$$\sigma(\varepsilon_3)=(0,\ 0,\ 0)=0\cdot\varepsilon_1+0\cdot\varepsilon_2+0\cdot\varepsilon_3,$$

则有

$$\sigma(\varepsilon_1\ \varepsilon_2\ \varepsilon_3)=(\varepsilon_1\ \ \varepsilon_2\ \ \varepsilon_3)\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}.$$

所以σ在基底 ε_1，ε_2，ε_3 上的表示矩阵为

$$A=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}.$$

由

$$\tau(\varepsilon_1)=(-1,\ 0,\ 0)=-\varepsilon_1$$
$$\tau(\varepsilon_2)=(0,\ 1,\ 0)=\varepsilon_2$$
$$\tau(\varepsilon_3)=(0,\ 0,\ 1)=\varepsilon_3,$$

所以τ在 ε_1，ε_2，ε_3 上的表示矩阵为

$$B=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

由

$$\rho(\varepsilon_1)=(1+0,\ 0+0,\ 0+1)=(1,\ 0,\ 1)=\varepsilon_1+\varepsilon_3$$
$$\rho(\varepsilon_2)=(0+1,1+0,0+0)=(1,1,\ 0)=\varepsilon_1+\varepsilon_2$$
$$\rho(\varepsilon_3)=(0+0,\ 0+1,\ 1+0)=(0,\ 1,\ 1)=\varepsilon_2+\varepsilon_3$$

所以ρ在 ε_1，ε_2，ε_3 上的表示矩阵为

$$C=\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix}.$$

进而可得 $\sigma+\tau$ 在 ε_1，ε_2，ε_3 上的表示矩阵为

$$A+B=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}+\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}0&0&0\\0&2&0\\0&0&1\end{pmatrix},$$

$\tau\rho$ 的表示矩阵为

$$BC=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix}=\begin{pmatrix}-1&-1&0\\0&1&1\\1&0&1\end{pmatrix}.$$

例 2 $D_n[x]=[1,\ x,\ x^2,\ \cdots,\ x^{n-1}]$。求线性变换

$$d:\ f(x)\longmapsto f'(x)$$

在 $1,\ x,\ x^2,\ \cdots,\ x^{n-1}$ 上的表示矩阵。

解 由

$$d(1)=0$$
$$d(x)=1$$
$$d(x^2)=2x$$
$$\vdots$$
$$d(x^{n-1})=(n-1)x^{n-2}.$$

则有

$$d(1,x,x^2,\cdots,x^{n-1})=(1\ x\ x^2\ \cdots\ x^{n-1})\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&2&\cdots&0&0\\0&0&0&\cdots&0&0\\\cdots&&\cdots&&\cdots&\\0&0&0&\cdots&n-2&0\\0&0&0&\cdots&0&n-1\\0&0&0&\cdots&0&0\end{pmatrix}.$$

所以d的表示矩阵为

$$D=\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&2&\cdots&0&0\\0&0&0&\cdots&0&0\\\cdots&&\cdots&&\cdots&\\0&0&0&\cdots&n-2&0\\0&0&0&\cdots&0&n-1\\0&0&0&\cdots&0&0\end{pmatrix}.$$

例 3 求例 1 中线性变换ρ的一维不变子空间。

解 我们知道，求 ρ 的一维不变子空间就是对每一个特征值求出一个特征向量。而这完全归结为求其表示矩阵的特征根及特征向量。

ρ 的表示矩阵为

$$C=\begin{pmatrix}1&1&0\\0&1&1\\1&0&1\end{pmatrix}.$$

求出 C 的特征根．即有

$$|\lambda I-C|=\begin{vmatrix}\lambda-1&-1&0\\0&\lambda-1&-1\\-1&0&\lambda-1\end{vmatrix}=(\lambda-1)^3-1$$

$$=(\lambda-2)(\lambda^2-\lambda+1).$$

从而 C 只有一个实特征根 2．于是 C 的属于特征根 2 的特征向量恰是齐次线性方程组

$$\begin{pmatrix}1&-1&0\\0&1&-1\\-1&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

的非零解．

容易看出上述方程组系数阵的秩等于 2，从而解空间是一维的．所以这个解空间就是 ρ 的一维不变子空间．于是，任意求出一个非零解，如 $\xi=(1, 1, 1)$ 就是一个非零解，这样 $S_\rho(2)=L(\xi)$ 就是 ρ 的一维不变子空间．

练　习　五

1．设 σ 是 $D^{(3)}$ 上的一个变换，如果

$\sigma(x_1, x_2, x_3)=(2x_1-x_2, x_2+x_3, x_1)$

证明：1) σ 是线性变换

2) 求 σ 在基底 $\varepsilon_1=(1, 0, 0)$，$\varepsilon_2=(0, 1, 0)$，$\varepsilon_3=(0, 0, 1)$ 上的表示矩阵．

2．设 σ 是 3 维线性空间 V 上的一个线性变换，在基底 ε_1，ε_2，

ε_3 上的表示矩阵为

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

1）求σ在基底 ε_3，ε_2，ε_1 上的表示矩阵。

2）求σ在基底 ε_1，$k\varepsilon_2$，ε_3 上的表示矩阵，其中 $k\in F$，且 $k\neq0$。

3）求σ在基底 $\varepsilon_1+\varepsilon_2$，$\varepsilon_2$，$\varepsilon_3$ 上的表示矩阵。

3. 设σ是 $F_n[x]$ 上的一个线性变换，如果

$$\sigma(f(x))=f'(x)$$

求σ在以下两个基底上的表示矩阵。

1) 1，$1+x$，$1+x^2$，$\cdots$，$1+x^{n-1}$

2) 1，$\dfrac{x-c}{1!}$，$\dfrac{(x-c)^2}{2!}$，$\cdots$，$\dfrac{(x-c)^{n-1}}{(n-1)!}$

4. 在 $D^{(3)}$ 中，试求在基底 $\varepsilon_1=(1,\ 0,\ 0)$，$\varepsilon_2=(1,\ 1,\ 0)$ $\varepsilon_3=(1,\ 1,\ 1)$ 上的表示矩阵为

$$A=\begin{pmatrix} 1 & -1 & 2 \\ -1 & 0 & -1 \\ 1 & 2 & 2 \end{pmatrix}$$

的线性变换 σ，并求向量 $\alpha=(1,\ 2,\ 3)$ 的象 $\sigma(\alpha)$。

5. 在$D^{(3)}$中，设基底 $\eta_1=(-1,\ 0,\ 2)$，$\eta_2=(0,\ 1,\ 1)$，$\eta_3=(3,\ -1,\ 0)$ 在线性变换σ下的象为

$$\sigma(\eta_1)=(-5,\ 0,\ 3)$$

$$\sigma(\eta_2)=(0,\ -1,\ 6)$$

$$\sigma(\eta_3)=(-5,\ -1,\ 9)$$

求σ在 η_1，η_2，η_3 上的表示矩阵。

§6　表示矩阵的变换公式

在上一节我们已充分地看到，对于有限维空间的线性变换，表示矩阵的重要作用．十分明白，线性变换的表示矩阵是对于事先取定的基底而言的．因此，对于给定的一个线性变换 σ，在空间V的每一个基底上都有一个表示矩阵．于是，一般来说，一个线性变换 σ 就可能有许多不同的表示矩阵．这样自然提出一个问题：同一个线性变换 σ 在空间V的各个基底上所有的表示矩阵之间有何联系？本节就来讨论这个问题．

设σ是n维线性空间V的线性变换．

$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是V的一个基底，σ 在这个基底上的表示矩阵为A；

$\delta_1, \delta_2, \cdots, \delta_n$ 也是V的一个基底，σ 在这个基底上的表示矩阵为B．

令从基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\delta_1, \delta_2, \cdots, \delta_n$ 的过渡阵为P，即有

$$(\delta_1, \delta_2, \cdots, \delta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)P. \tag{1}$$

于是则有

$$\sigma(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A \tag{2}$$

$$\sigma(\delta_1, \delta_2, \cdots, \delta_n) = (\delta_1, \delta_2, \cdots, \delta_n)B \tag{3}$$

将（1）式代入（3），得

$$\sigma(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)P = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)PB$$

$$\sigma(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)PBP^{-1} \tag{4}$$

比较（2）与（4）两个式子，由线性变换在同一个基底上的表示矩阵是唯一的，便得

$$A = PBP^{-1} \text{ 即 } B = P^{-1}AP.$$

这样，我们已经证明了

定理 1　**设σ是n维线性空间V的线性变换．**$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ **与**

δ_1，δ_2，$\cdots$，δ_n 是V的两个基底，过渡阵为P。如果σ在基底 ε_1，ε_2，$\cdots$，ε_n 上的表示矩阵为A，在基底δ_1，δ_2，$\cdots$，δ_n 上的表示矩阵为B，那么 $B=P^{-1}AP$。

这个定理表明，同一个线性变换 σ 在空间V的各个基底上的所有表示矩阵都是彼此相似的，而且演化阵就是基底的过渡阵。进而，我们提出这样的问题：如果A是σ在基底 ε_1，ε_2，$\cdots$，ε_n 上的表示矩阵，B是与A相似的任一矩阵，即 $B=P^{-1}AP$。问B是否也是 σ 在某一基底上的表示矩阵？从定理 1 的推证过程已明显看出解答这个问题的线索。如果有某一个基底 δ_1，δ_2，$\cdots$，δ_n 使σ在此基底上的表示矩阵为B的话，那么从基底 ε_1，ε_2，$\cdots$，ε_n 到基底δ_1，δ_2，$\cdots$，δ_n 的过渡阵必定是演化阵P。因此，我们以从A到B的演化阵P为过渡阵得一基底 δ_1，δ_2，$\cdots$，δ_n：

$$(\delta_1\ \delta_2\ \cdots\delta_n)=(\varepsilon_1\ \varepsilon_2\cdots\varepsilon_n)P,$$

于是，由定理 1 便知，σ 在基底 δ_1，δ_2，$\cdots$，δ_n 上的表示矩阵恰好就是B。

这样，我们又证明了

定理 2　若σ在 ε_1，ε_2，$\cdots$，ε_n 上的表示矩阵为 A，B 是相似于A的任一矩阵：$B=P^{-1}AP$。那么以P为过渡阵得到的基底 δ_1，δ_2，$\cdots$，δ_n，使σ在此基底上的表示矩阵为B。

把以上两个定理概括起来说就是一句话：同一个线性变换所有的表示矩阵恰好是一个相似类。这一事实使我们更加明确的认识到，n 阶方阵A的任何一种在相似关系之下不变的性质都体现着它所表示的那个线性变换σ的一种特性。比如，一类相似矩阵的特征多项式是唯一确定的。正象我们在上一节已经知道的那样，这个特征多项式的一切根恰是线性变换 σ 的全部特征值。这样，自然可以把σ的表示矩阵A的特征多项式也叫做σ的特征多项式，从而σ 的特征值也叫做σ的特征根。

由于线性变换与其表示矩阵之间特定的对应关系，自然可得：如果 $f(\lambda)$ 是σ 的特征多项式，那么 $f(\sigma)=0$。这样，我们又可给

出σ的最小多项式的概念，而且容易知道，σ的最小多项式与其表示矩阵A的最小多项式是一致的。

例1　在$D^{(3)}$中，取基底为

$$\varepsilon_1=(1,\ 0,\ 0),\ \varepsilon_2=(0,\ 1,\ 0),\ \varepsilon_3=(0,\ 0,\ 1)$$

已知线性变换

$$\sigma:\ (x,\ y,\ z)\longmapsto(x,\ y,\ 0)$$

在 ε_1，ε_2，ε_3 上的表示矩阵为

$$A=\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}$$

现在另取$D^{(3)}$ 的一个基底

$$\delta_1=(1,\ 0,\ 0),\ \delta_2=(1,\ 1,\ 0),\ \delta_3=(1,\ 1,\ 1).$$

求σ在 δ_1，δ_2，δ_3 上的表示矩阵B。

解　先找到从 ε_1，ε_2，ε_3 到δ_1，δ_2，δ_3的过渡阵P：

$$(\delta_1\ \delta_2\ \delta_3)=(\varepsilon_1\ \varepsilon_2\ \varepsilon_3)\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}.$$

所以

$$P=\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}.$$

再算出P^{-1}：

$$P^{-1}=\begin{pmatrix}1&-1&0\\0&1&-1\\0&0&1\end{pmatrix}.$$

最后便得

$$B=P^{-1}AP=\begin{pmatrix}1&-1&0\\0&1&-1\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&1&1\\0&1&1\\0&0&1\end{pmatrix}=$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

这里按表示矩阵的定义亦可求出 B：

$\sigma(\bar{\delta}_1)=\bar{\delta}_1$

$\sigma(\bar{\delta}_2)=\bar{\delta}_2$

$\sigma(\bar{\delta}_3)=\bar{\delta}_2$

从而σ在 $\bar{\delta}_1$，$\bar{\delta}_2$，$\bar{\delta}_3$ 上的表示矩阵

$$B=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

例 2 在$D_4[x]$ 中，取定基底为 1，x，x^2，x^3. 线性变换

$$d:\ f(x)\longmapsto f'(x)$$

在 1，x，x^2，x^3 上的表示矩阵

$$D=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

给一可逆阵

$$P=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

以 P 为演化矩阵得相似于 D 的阵

$$C=P^{-1}DP=\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}=$$

$$=\begin{pmatrix} 0 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix},$$

求$D_4[x]$ 的一个基底 δ_1，δ_2，δ_3，δ_4 使d的表示矩阵为 C。

解　按定理2，此基底就是以P为过渡阵从 1，x，x^2，x^3 得到的，即

$$(\delta_1\ \delta_2\ \delta_3\ \delta_4) = (1\ x\ x^2\ x^3)P = (x\ 1\ x^3\ x^2)$$

例3　在$D_4[x]$ 中，已知求导数的变换d有以下的表示矩阵

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

于是容易计算

$$D^2 = \begin{pmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\quad D^3 = \begin{pmatrix} 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},\quad D^4 = 0.$$

这说明 $f(\lambda) = \lambda^4$ 是D的特征多项式同时也是最小多项式。于是按定义，线性变换d的特征多项式是 $f(\lambda) = \lambda^4$，最小多项式也是$f(\lambda) = \lambda^4$。

练　习　六

1．设σ是三维线性空间V上的线性变换，在基底α_1，α_2，α_3 上的表示矩阵为

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 2 & 0 \end{pmatrix}$$

求σ在基底 $\beta_1 = -\alpha_1 + \alpha_3$，$\beta_2 = \alpha_1 + \alpha_2$，$\beta_3 = -\alpha_1 - \alpha_2 + \alpha_3$ 上的表示矩阵。

2．设$D^{(3)}$的两个基底为

$\varepsilon_1 = (1,\ 0,\ 1)$，$\varepsilon_2 = (2,\ 1,\ 0)$，$\varepsilon_3 = (1,\ 1,\ 1)$

$\eta_1 = (1,\ 2,\ -1)$，$\eta_2 = (2,\ 2,\ -1)$，$\eta_3 = (2,\ -1,\ -1)$

如果线性变换σ使得

$\sigma(\varepsilon_i) = \eta_i$，$i = 1,\ 2,\ 3$。

1）求由 ε_1，ε_2，ε_3 到 η_1，η_2，η_3 的过渡阵，

2）求σ在 ε_1，ε_2，ε_3 上的表示矩阵，

3）求σ在 η_1，η_2，η_3上的表示矩阵。

§7 线性变换的标准表示矩阵

我们已知，n维线性空间V的每一线性变换σ，在V的任意一个基底上都有一个表示矩阵，并且σ在各个基底上的一切表示矩阵恰好组成一个相似类。这一事实对于用矩阵来表示线性变换似乎带来一些麻烦。但正是由于一个线性变换有许多表示矩阵，才使得用矩阵表示线性变换有了选择的可能。这样，把n阶方阵的相似标准形理论用于线性变换的表示矩阵，即有

定理1 **设V是F上n维线性空间，σ是V上任一线性变换。那么存在V的基底 α_1，α_2，…，α_n 使σ在这个基底上的表示矩阵为有理标准形。**

定理2 **设V是复数域C上n维线性空间，σ是V上任一线性变换。那么存在V的基底β_1，β_2，…，β_n 使σ在这个基底上的表示矩阵为若当标准形。**

例1 在 $D_4[x]$ 中考虑求导数的变换 d。找出 $D_4[x]$ 的基底 α_1，α_2，α_3，α_4 使d在这个基底上的表示矩阵为有理标准形。

解 先任取 $D_4[x]$ 的一个基底，如 1，x，x^2，x^3。已知d在这个基底上的表示矩阵为

$$A=\begin{pmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\\0&0&0&0\end{pmatrix}.$$

其次求四阶方阵A的有理标准形，为此需要计算A的不变因子。即求A的特征矩阵

$$\lambda I - A = \begin{pmatrix} \lambda & -1 & 0 & 0 \\ 0 & \lambda & -2 & 0 \\ 0 & 0 & \lambda & -3 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

的标准形，其对角线上的非零元素就是A的不变因子．一般可用初等变换把特征矩阵化简成标准形，从而得到所求的不变因子，或者通过计算行列式因子，利用行列式因子与不变因子之间的关系，从而求出不变因子．对于现在这个特征矩阵来说，它的行列式因子比较简单，因此用计算行列式因子的办法来求不变因子比较方便．事实上，由于 $\lambda I - A$ 有一个三阶子式等于非零常数，所以它的三阶行列式因子 $D_3(\lambda) = 1$． 从而 $D_1(\lambda) = 1$， $D_2(\lambda) = 1$． 另外显然有 $D_4(\lambda) = \lambda^4$．这样由这四个行列式因子马上就得出不变因子为

$$d_1(\lambda) = 1,\ d_2(\lambda) = 1,\ d_3(\lambda) = 1,\ d_4(\lambda) = \lambda^4.$$

于是A的有理标准形为

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

不难想到，d 的这个表示矩阵所依赖的基底是

$$\alpha_1 = \frac{x^3}{3!},\quad \alpha_2 = \frac{x^2}{2!},\quad \alpha_3 = \frac{x}{1!},\quad \alpha_4 = 1.$$

例 2 $F^{(3)}$ 的线性变换

$$\sigma:\ (x,\ y,\ z) \longrightarrow (z,\ y,\ x)$$

求σ的有理标准形的和若当标准形的表示矩阵．

解 先任取$F^{(3)}$ 的一个基底，如

$$\varepsilon_1 = (1,\ 0,\ 0),\ \varepsilon_2 = (0,\ 1,\ 0),\ \varepsilon_3 = (0,\ 0,\ 1).$$

于是σ在这个基底上的表示矩阵为

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

其次求A的不变因子、初等因子．即有

$$\lambda I - A = \begin{pmatrix} \lambda & 0 & -1 \\ 0 & \lambda-1 & 0 \\ -1 & 0 & \lambda \end{pmatrix} \longrightarrow \begin{pmatrix} -1 & 0 & \lambda \\ 0 & \lambda-1 & 0 \\ \lambda & 0 & -1 \end{pmatrix} \longrightarrow$$

$$\begin{pmatrix} -1 & 0 & \lambda \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda^2-1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda^2-1 \end{pmatrix}.$$

因而$\mathscr{A}$的不变因子为

$$d_1(\lambda) = 1, \ d_2(\lambda) = \lambda - 1, \ d_3(\lambda) = \lambda^2 - 1,$$

初等因子为

$$\lambda - 1, \ \lambda - 1, \ \lambda + 1.$$

这样便得$\mathscr{A}$的有理标准形为

$$B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

若当标准形为

$$B_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

换句话说就是，B_1与B_2分别是线性变换σ的有理标准形的表示矩阵和若当标准形的表示矩阵。

习 题 十二

1．证明：若 $\alpha_1, \alpha_2, \cdots, \alpha_m$ $(m \leqslant n)$ 是 $V_n(F)$ 的线性无关向量组，$\beta_1, \beta_2, \cdots, \beta_m$ 是$V_n(F)$ 中任意一组向量，则在$V_n(F)$上存在线性变换 σ，使得

$$\sigma(\alpha_i) = \beta_i \qquad i = 1, \ 2, \ \cdots, \ m.$$

2．求下列线性变换在所指定基底上的表示矩阵。

1）在 $F_n[x]$ 中，设线性变换σ为

$$\sigma(f(x))=f(x+1)-f(x)$$

基底为

$$\varepsilon_0=1,\qquad \varepsilon_i=\frac{x(x-1)\cdots(x-i+1)}{i!};\quad i=1,\ 2,\ \cdots,\ n-1$$

2）设六个函数

$$\varepsilon_1=e^{ax}\cos bx\qquad \varepsilon_2=e^{ax}\sin bx$$

$$\varepsilon_3=xe^{ax}\cos bx\qquad \varepsilon_4=xe^{ax}\sin bx$$

$$\varepsilon_5=\frac{1}{2}x^2e^{ax}\cos bx\qquad \varepsilon_6=\frac{1}{2}x^2e^{ax}\sin bx$$

求由它们生成的线性空间 $L(\varepsilon_1,\ \varepsilon_2,\ \varepsilon_3,\ \varepsilon_4,\ \varepsilon_5,\ \varepsilon_6)$ 中线性变换

$$\sigma(f(x))=f'(x)$$

在该基底上的表示矩阵

3．在 $M_2(F)$ 中定义线性变换

$$\sigma_1(X)=\begin{pmatrix}a & b\\ c & d\end{pmatrix}X$$

$$\sigma_2(X)=X\begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

$$\sigma_3(X)=\begin{pmatrix}a & b\\ c & d\end{pmatrix}X\begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

求σ_1，σ_2，σ_3 在基底

$$E_{11}=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\ E_{12}=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\ E_{21}=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},\ E_{22}=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

上的表示矩阵。

4．在n 维线性空间V中，证明V中与全体线性变换可交换的线性变换是相似变换。

5．设σ是线性空间 $V_n(F)$ 中的线性变换，证明如果σ在任意一组基底上的表示矩阵都相同，则σ是相似变换。

6．设α_1，α_2 是线性变换σ属于不同特征根 λ_1，λ_2 的特征向量，证明 $\alpha_1+\alpha_2$ 不是σ的特征向量。

7．线性变换是相似变换的充分必要条件是线性空间V中每个非零向量都是它的特征向量．

8．在 $F_n[x]$ 中 $(n>1)$，设线性变换σ为

$$\sigma(f(x))=f'(x)$$

证明σ在任何一组基底上的表示矩阵，都不可能是对角阵。

9．设复数域上线性空间V的线性变换σ，τ，如果$\sigma\tau=\tau\sigma$，证明

1）如果λ_0是σ的特征根，则$S_\sigma(\lambda_0)$是τ的不变子空间。

2）σ，τ 至少有一个共公的特征向量。

10．设V是复数域C上的n维线性空间，而线性变换 σ 在基底 ε_1，ε_2，…，ε_n 上的表示矩阵为一若当矩阵，证明：

1）V中包含 ε_1 的不变子空间只有V自己，

2）V中任一非零不变子空间都包含 ε_n，

3）V不能分解成两个非平凡子空间的直和．

11．由n维线性空间V的全体线性变换组成的线性空间是n^2维的。

12．设σ是 $V_n(F)$ 的一个线性变换，证明：

1）在 $F[x]$ 中有一个次数$\leqslant n^2$的多项式 $f(x)$，使得

$f(\sigma)=0$，

2）设 $(f(x), g(x))=d(x)$，如果 $f(\sigma)=0$，$g(\sigma)=0$ 证明：$d(\sigma)=0$，

3）σ 是可逆线性变换的充分必要条件是有一常数项不为 0 的多项式 $f(x)$ 使得

$$f(\sigma)=0$$

13．设σ是线性空间V上的可逆线性变换，证明

1）σ 的特征根一定不为 0 。

2）如果λ是σ的特征根，则λ^{-1}是σ^{-1}的特征根。

14．设 σ 是线性空间 $V_n(F)$ 上的线性变换，证明σ的表示矩阵的行列式 $|A|=0$ 的充分必要条件是σ以 0 为特征根。

15. 设 σ_1, σ_2, …, σ_s 是线性空间V上的两两不同的线性变换，那么V中必存在一个向量α使得 $\sigma_1(\alpha)$, $\sigma_2(\alpha)$, …, $\sigma_s(\alpha)$ 也两两不同.

16. 设 σ, τ是线性空间V上的线性变换，如果 $\sigma^2=\sigma$, $\tau^2=\tau$, 证明

1) σ与τ有相同的值域的充分必要条件是 $\sigma\tau=\tau$, $\tau\sigma=\sigma$

2) σ 与τ有相同的核的充分必要条件是: $\sigma\tau=\sigma$, $\tau\sigma=\tau$.

17. 设 ε_1, ε_2, …, ε_n 是V的基底, A, B 是n 阶方阵，如果

$$(\alpha_1, \alpha_2, \cdots, \alpha_n)=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A \tag{1}$$

$$(\beta_1, \beta_2, \cdots, \beta_n)=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)B \tag{2}$$

当 α_1, α_2, …, α_n 是线性无关，且σ是V中的线性变换，使得

$$\sigma(\alpha_i)=\beta_i \tag{3}$$

求σ在 ε_1, ε_2, …, ε_n 上的表示矩阵.

18. 设σ是 $V_n(F)$ 上的线性变换，W_1, W_2 是 $V_n(F)$ 的子空间，如果

$$V_n(F)=W_1\oplus W_2$$

证明σ是可逆线性变换的充分必要条件是

$$V_n(F)=\sigma(W_1)\oplus\sigma(W_2)$$

19. 设 S 是$V_n(F)$ 中一些线性变换的集合，W是$V_n(F)$ 的子空间，如果W对 S 中每个线性变换都是不变的，则称W是 S 的不变子空间，如果 S 没有非平凡的不变子空间，则称 S 是不可约的. 设σ是 $V_n(F)$ 的一个线性变换， 如果它与不可约的 S 中每个线性变换可交换，证明σ或者是零变换，或者是可逆线性变换.

20. 设V是一个n维线性空间，σ是V的一个线性变换，则必存在自然数 $k\leqslant n$, 使得

$$\sigma^kV=\sigma^{k+1}V=\cdots.$$

第十三章　欧氏空间及其线性变换

§1　欧氏空间的定义及简单性质

在线性空间的基本的直观的模型，通常的几何空间中，除了线性运算之外，还有大家熟知的主要运算——内积运算．通过内积便能表述基本的度量性质：向量的长，两个向量间的夹角等．一般线性空间的度量性质在许多问题中有着特殊的地位，因此有必要对一般线性空间引进内积运算，进而讨论它的基本度量性质．

定义1　设V是实数域D上的线性空间，在V中规定一个叫做内积的运算，即对V中任二向量α，β其内积是D中的一个实数，记作(α,β)．如果具有以下性质

1）$(\alpha,\beta)=(\beta,\alpha)$

2）$(k\alpha,\beta)=k(\alpha,\beta)$

3）$(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$

4）$(\alpha,\alpha)\geqslant 0$　只有$\alpha=0$时$(\alpha,\alpha)=0$

这里α，β，γ是V中任意向量，k为任意实数．则称V是欧几里得空间．简称欧氏空间．

例1　在二，三维几何空间$D^{(2)}$，$D^{(3)}$中，分别规定内积为

$$(\alpha,\beta)=a_1b_1+a_2b_2\qquad \alpha=(a_1,a_2),\ \beta=(b_1,b_2)$$

$$(\xi,\eta)=a_1b_1+a_2b_2+a_3b_3$$

其中$\xi=(a_1,\ a_2,\ a_3)$，$\eta=(b_1,\ b_2,\ b_3)$，那么$D^{(2)}$，$D^{(3)}$都是欧氏空间．

例2　在n维线性空间$D^{(n)}$中，规定内积

$$(\alpha, \ \beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

其中 $\alpha = (a_1, \ a_2, \cdots, \ a_n)$，$\beta = (b_1, \ b_2, \ \cdots, \ b_n)$，则 $D^{(n)}$ 做成欧氏空间。

例 3　在连续函数空间 $C[a, b]$ 中，规定内积为

$$(f(x), \ g(x)) = \int_a^b f(x) g(x) dx$$

则 $C[a, b]$ 做成欧氏空间。

内积的四条性质可验证如下：

1）因为

$$\int_a^b f(x) g(x) dx = \int_a^b g(x) f(x) dx$$

所以

$$(f(x), \ g(x)) = (g(x), \ f(x))$$

2）$\int_a^b kf(x), g(x) dx = k\int_a^b f(x) \cdot g(x) dx$

所以

$$(kf(x), g(x)) = k(f(x), \ g(x))$$

3）$\int_a^b (f(x) + h(x)) \ g(x) dx = \int_a^b f(x) g(x) dx + \int_a^b h(x) \cdot g(x) dx$

所以

$$(f(x) + h(x), \ g(x)) = (f(x), \ g(x)) + (h(x), \ g(x))$$

4）$\int_a^b f(x) \ f(x) dx \geqslant 0$ 且只在 $f(x) = 0$ 时才有 $\int_a^b f^2(x) dx = 0$ 所以

$(f(x), \ f(x)) \geqslant 0$, 且只在 $f(x) = 0$ 时，才有

$(f(x), \ f(x)) = 0$。

如果 V 是欧氏空间，S 是 V 的线性子空间．这时用 V 的内积运算，S 本身也是欧氏空间．今后简称 S 为 V 的子空间．

如 $S=\{(a,\ b,\ 0)\in D^{(3)}\,|a,b\in D\}$，是 $D^{(3)}$ 的做为欧氏空间的子空间。

由定义 1 容易证明

命题 1　在欧氏空间中，下列诸等式成立：

1° $(\alpha,k\beta)=k(\alpha,\beta)$

2° $(\alpha,\beta+\gamma)=(\alpha,\beta)+(\alpha,\gamma)$

3° $(\alpha,\ 0)=(0,\alpha)=0$

4° $(k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r,\ h_1\beta_1+h_2\beta_2+\cdots+h_s\beta_s)$

$$=\sum_{i,j=1}^{r,s}k_ih_j(\alpha_i,\ \beta_j).$$

事实上

$$(\alpha,\ k\beta)=(k\beta,\alpha)=k(\beta,\alpha)=k(\alpha,\beta)$$

$$(\alpha,\ \beta+\gamma)=(\beta+\gamma,\alpha)=(\beta,\ \alpha)+(\gamma,\ \alpha)=(\alpha,\beta)+(\alpha,\gamma)$$

$$(\alpha,0)=(\alpha,0\cdot 0)=0(\alpha,0)=0$$

$$\begin{aligned}&(k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r,\ h_1\beta_1+h_2\beta_2+\cdots+h_s\beta_s)\\&=(k_1\alpha_1,\ h_1\beta_1+h_2\beta_2+\cdots+h_s\beta_s)+\cdots+(k_r\alpha_r,\ h_1\beta_1+\cdots\\&\quad+h_s\beta_s)=(k_1\alpha_1,h_1\beta_1)+(k_1\alpha_1,h_2\beta_2)+\cdots+(k_1\alpha_1,h_s\beta_s)+\\&\quad\cdots+(k_r\alpha_r,\ h_1\beta_1)+(k_r\alpha_r,\ h_2\beta_2)+\cdots+(k_r\alpha_r,\ h_s\beta_s)\\&=k_1h_1(\alpha_1,\beta_1)+k_1h_2(\alpha_1,\ \beta_2)+\cdots+k_1h_s(\alpha_1,\ \beta_s)+\cdots\\&\quad+k_rh_1(\alpha_r,\ \beta_1)+k_rh_2(\alpha_r,\ \beta_2)+\cdots+k_rh_s(\alpha_r,\beta_s)\\&=\sum_{i,j=1}^{r,s}k_ih_j(\alpha_i,\ \beta_j)\end{aligned}$$

下面我们利用内积给出向量长度，两个向量的交角的概念。

定义 2　设V是欧氏空间，$\alpha\in V$，称非负实数$\sqrt{(\alpha,\alpha)}$为向量α的长，记作$|\alpha|$。

定义 3　设V是欧氏空间，α，$\beta\in V$，当且仅当 $(\alpha,\beta)=0$时，称 α，β 正交，记作 $\alpha\perp\beta$。

如果 $|\alpha|=1$，则称 α 为**单位向量**或**标准向量**。

如果一组向量

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_s \tag{1}$$

$|\alpha_1|=|\alpha_2|=\cdots=|\alpha_s|=1$，则称（1）是**标准向量组**；如果 $i\neq j$ 时，$\alpha_i\perp\alpha_j$，则称（1）是**正交向量组**，特别地 $s=1$ 时，也说（1）是正交向量组。当（1）是标准向量组，又是正交向量组时，就说（1）是**标准正交组**。

容易指出，如果 $\alpha\neq 0$，则

$$\beta=\frac{\alpha}{|\alpha|}$$

是单位向量。如果 $\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_s$ 是正交组，则 $k_1\alpha_1,\ k_2\alpha_2,\ \cdots,\ k_s\alpha_s$ 也是正交组。特别地，$\frac{\alpha_1}{|\alpha_1|},\ \frac{\alpha_2}{|\alpha_2|},\ \cdots,\ \frac{\alpha_s}{|\alpha_s|}$ 是标准正交组。显然，过去对 n 数组空间规定的向量正交性概念与做为欧氏空间来看的 $D^{(n)}$ 中的向量正交性概念是一致的。

命题 2 $|k\alpha|=|k|\,|\alpha|$

事实上，$|k\alpha|=\sqrt{(k\alpha,\ k\alpha)}=\sqrt{k^2(\alpha,\alpha)}=|k|\sqrt{(\alpha,\alpha)}$
$=|k|\,|\alpha|$.

命题 3 设 $\alpha\in V$，那么对任意 $\xi\in V$ 都有

$(\alpha,\xi)=0$ 当且仅当 $\alpha=0$.

事实上，充分性是自明的，取 $\xi=\alpha$ 时，即可说明必要性是正确的。

命题 4 若 $(\beta,\alpha_i)=0$，$(i=1,\ 2,\ \cdots,\ s)$ 则

$$\left(\beta,\sum_{i=1}^{s}k_i\alpha_i\right)=0.$$

因为 $\left(\beta,\ \sum_{i=1}^{s}k_i\alpha_i\right)=\sum_{i=1}^{s}k_i(\beta,\ \alpha_i)=0$。

推论　设 S 为 V 的一个子集，

$$W=\{\xi \mid (\alpha,\xi)=0,\ \forall \alpha\in S\}$$

那么 W 是 V 的子空间，记作 $W=S^{\perp}$。

为了规定夹角的概念我们需要重要的不等式，即

命题 5　对欧氏空间 V 中任二向量 α，β 有

$$|(\alpha,\beta)|\leqslant|\alpha|\ |\beta| \qquad (2)$$

而等号成立当且仅当 α，β 是线性相关的。

证　当 α，β 线性相关时，令 $\alpha=k\beta$，于是

$$|(\alpha,\beta)|=|(k\beta,\ \beta)|=|k|\ |(\beta,\beta)|=|k|\ (\beta,\beta)$$

而

$$|\alpha|\ |\beta|=|k\beta|\ |\beta|=|k|\ |\beta|\ |\beta|=|k|\ (\beta,\beta)$$

即有

$$|(\alpha,\beta)|=|\alpha|\ |\beta|.$$

当 α，β 线性无关时，对任意实数 t，有

$$\alpha+t\beta\neq0$$

于是

$$(\alpha+t\beta,\ \alpha+t\beta)>0$$

取　$t=-\dfrac{(\alpha,\beta)}{(\beta,\beta)}$，则有

$$(\alpha,\ \alpha)-\frac{(\alpha,\beta)^2}{(\beta,\beta)}>0$$

从而

$$(\alpha,\beta)^2<(\alpha,\alpha)\cdot(\beta,\beta)$$

两边开方即得

$$|(\alpha,\beta)|<|\alpha|\ |\beta|$$

一般称（2）为柯西不等式。

推论　$|\alpha+\beta|\leqslant|\alpha|+|\beta|$.

$$\begin{aligned}|\alpha+\beta|^2&=(\alpha+\beta,\ \alpha+\beta)=(\alpha,\alpha)+2(\alpha,\beta)+(\beta,\beta)\\&\leqslant|\alpha|^2+2|\alpha|\cdot|\beta|+|\beta|^2\end{aligned}$$

$$=(|\alpha|+|\beta|)^2$$

所以

$$|\alpha+\beta|\leqslant|\alpha|+|\beta|$$

由命题5对欧氏空间的任二非零向量 α，β，有

$$\frac{|(\alpha,\beta)|}{|\alpha|\ |\beta|}\leqslant 1$$

即

$$-1\leqslant\frac{(\alpha,\beta)}{|\alpha|\ |\beta|}\leqslant 1$$

这样0与 π 之间恰有一个角度θ，使

$$\cos\theta=\frac{(\alpha,\beta)}{|\alpha|\ |\beta|}$$

即

$$\theta=\cos^{-1}\frac{(\alpha,\beta)}{|\alpha|\ |\beta|}$$

我们称角 θ 为 α 与 β 的夹角，记作 $\langle\alpha,\beta\rangle$。于是 $\alpha\perp\beta$ 当且仅当 $\langle\alpha,\beta\rangle=\frac{\pi}{2}$。

把柯西不等式用于 $D^{(n)}$，$C[a,b]$ 这两个具体的空间上就得到两个常见的重要不等式。

$$|a_1b_1+a_2b_2+\cdots+a_nb_n|\leqslant\sqrt{a_1^2+a_2^2+\cdots+a_n^2}\cdot\sqrt{b_1^2+b_2^2+\cdots+b_n^2}$$

$$\left|\int_a^b f(x)\cdot g(x)\,dx\right|\leqslant\left(\int_a^b f^2(x)\,dx\right)^{\frac{1}{2}}\cdot\left(\int_a^b g^2(x)\,dx\right)^{\frac{1}{2}}.$$

命题6 若 α_1，α_2，…，α_s 是一个正交组，当中没有零向量，那么 α_1，α_2，…，α_s 必线性无关。

证 考查

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$$

于是

$$(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s,\ \alpha_1) = 0$$

$$k_1(\alpha_1, \alpha_1) + k_2(\alpha_2, \alpha_1) + \cdots + k_s(\alpha_s, \alpha_1) = 0$$

从而

$$k_1(\alpha_1, \alpha_1) = 0$$

因 $\alpha_1 \neq 0$，所以 $k_1 = 0$.

同理可得 $k_2 = 0$，…，$k_s = 0$。故 α_1，$\alpha_2, \cdots$，α_s 线性无关。

推论 设 W_1，$W_2, \cdots$，W_s 是一组两两正交的子空间，即当 $i \neq j$ 时，对任意 $\alpha_i \in W_i$，$\alpha_j \in W_j$，$(\alpha_i, \alpha_j) = 0$，则和 $W_1 + W_2 + \cdots + W_s$ 是直和。

事实上，由 $\alpha_1 + \alpha_2 + \cdots + \alpha_s = 0$，必有

$$\alpha_1 = 0,\ \alpha_2 = 0,\ \cdots,\ \alpha_s = 0$$

故知和 $W_1 + W_2 + \cdots + W_s$ 是直和。

练 习 一

1. 在 $D^{(3)}$ 中规定数积如下：

1）$(\alpha, \beta) = a_1b_3 + a_2b_2 + a_3b_1$

2）$(\alpha, \beta) = a_1b_1 + a_2b_2 + a_3b_3 + \frac{1}{2}a_1b_3 + \frac{1}{2}a_3b_1$

其中 $\alpha = (a_1,\ a_2,\ a_3)$，$\beta = (b_1,\ b_2,\ b_3)$， 试判断 $D^{(3)}$ 按哪种数积构成欧氏空间。

2. 在欧氏空间里，对任意向量 α，β，证明以下等式：

1）$|\alpha+\beta|^2 + |\alpha-\beta|^2 = 2(|\alpha|^2 + |\beta|^2)$

2）$(\alpha,\ \beta) = \frac{1}{4}|\alpha+\beta|^2 - \frac{1}{4}|\alpha-\beta|^2$

3. 在欧氏空间里，若 α，β 正交，则

$$|\alpha+\beta|^2 = |\alpha|^2 + |\beta|^2$$

4. 在 $D^{(4)}$ 中，按通常的内积，求一单位向量与向量 $\alpha = (1, 1, -1, 1)$，$\beta = (1, -1, -1, 1)$，$\gamma = (2, 1, 1, 3)$ 正交。

5. 设 α, β 是欧氏空间中任二非零向量，证明：

1）$\alpha=k\beta$, $(k>0)$ 当且仅当 α, β 之间的夹角为 0，

2）$\alpha=k\beta$ $(k<0)$ 当且仅当 α, β 之间的夹角为 π.

6. 在欧氏空间 $D^{(n)}$ 中，求向量 $\alpha=(1,1,\cdots,1)$ 与每一向量

$\varepsilon_i=(0,\cdots,0,\overset{(i)}{1},0,\cdots,0)$, $i=1,2,\cdots,n$, 的夹角。

7. 证明：对任意实数 a_1, a_2, $\cdots$, a_n 有

$$\sum_{i=1}^{n}|a_i|\leqslant\sqrt{n(a_1^2+a_2^2+\cdots+a_n^2)}.$$

8. 设 V 是 n 维欧氏空间，证明：

1）如果 W 是 V 的子空间，那么

$(W^{\perp})^{\perp}=W$.

2）如果 W_1, W_2 是 V 的子空间，且 $W_1\subseteq W_2$，那么

$W_2^{\perp}\subseteq W_1^{\perp}$.

9. 设 α_1, α_2, $\cdots$, α_n 是欧氏空间 V 的一个基底，如果 $\gamma\in V$，且有

$$(\gamma,\alpha_i)=0$$

证明：$\gamma=0$

10. 设 V 是一欧氏空间，如果 α_1, $\alpha_2\in V$，且对任一 $\beta\in V$，都有

$$(\alpha_1,\beta)=(\alpha_2,\beta)$$

证明：$\alpha_1=\alpha_2$.

§2 有限维欧氏空间 标准正交基底

本节主要讨论有限维欧氏空间，指出从有限维线性空间构成欧氏空间的一般方法，通过特殊的基底——标准正交基底确定有限维欧氏空间的结构。

若 V 是欧氏空间，V 当然是线性空间，就把线性空间的维数叫

做欧氏空间的维数。设 $\dim V=n$，ε_1，ε_2，$\cdots$，ε_n 为任一基底，于是对任意向量 α，β

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$$
$$\beta=b_1\varepsilon_1+b_2\varepsilon_2+\cdots+b_n\varepsilon_n$$

则有

$$(\alpha,\ \beta)=\left(\sum_{i=1}^{n}a_i\varepsilon_i,\ \sum_{j=1}^{n}b_j\varepsilon_j\right)=\sum_{i,j=1}^{n}a_ib_j(\varepsilon_i,\ \varepsilon_j) \qquad (1)$$

（1）式表明，只要基底向量内积 $(\varepsilon_i,\ \varepsilon_j)$ 知道了，任二向量 α，β 的内积便可由其坐标 a_1，a_2，$\cdots$，a_n 与 b_1，b_2，$\cdots$，b_n 按（1）式来计算，令 $g_{ij}=(\varepsilon_i,\ \varepsilon_j)$，称$n$阶方阵

$$G=\begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1n}\\ g_{21} & g_{22} & \cdots & g_{2n}\\ \cdots & & \cdots & \\ g_{n1} & g_{n2} & \cdots & g_{nn}\end{pmatrix}$$

为基底 ε_1，ε_2，$\cdots$，ε_n **的度量矩阵**。于是在基底 ε_1，ε_2，$\cdots$，ε_n上内积计算公式（1）便可写成

$$(\alpha,\beta)=(a_1\ \ a_2\ \ \cdots\ \ a_n)\ \ G\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{pmatrix}$$

由 $(\varepsilon_i,\ \varepsilon_j)=(\varepsilon_j,\ \varepsilon_i)$ 可知 $G'=G$ 即度量矩阵是对称阵。

由 $(\xi,\xi)\geqslant 0$，只有当 $\xi=0$ 时才有 $(\xi,\xi)=0$，即

$$(x_1\ \ x_2\ \ \cdots\ \ x_n)\ \ G\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}\geqslant 0$$

只当 $x_1=x_2=\cdots=x_n=0$ 时才有

$$(x_1\ x_2\ \cdots\ x_n)\ G\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=0$$

其中 $\xi=(x_1, x_2, \cdots, x_n)$，可知度量矩阵$G$是正定的。总括言之，就是

对给定的n维欧氏空间V，任意取定一个基底$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$那么都确定一个 n 阶正定矩阵——基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$的度量矩阵：

$G=(g_{ij}), g_{ij}=(\varepsilon_i, \varepsilon_j)$

使V中内积运算的计算公式为

$$(\alpha,\beta)=(a_1\ a_2\cdots a_n)\ \ G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

其中 $\alpha=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$，$\beta=(\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$。

反之如果V是实数域D上的 n 维线性空间，任取基底$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，令 $G=(g_{ij})$ 为任意给定的正定矩阵。于是在V中规定内积为

$$(\alpha,\beta)=(a_1a_2\cdots\ a_n)\ \ G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}\qquad(2)$$

其中 $\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$，$\beta=b_1\varepsilon_1+b_2\varepsilon_2+\cdots+b_n\varepsilon_n$，那么则有

1） $(\alpha,\beta)=(\beta,\alpha)$

事实上，$(\alpha,\beta)=(a_1\ a_2\cdots a_n)\ G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$ 是一个实数，

所以

$$(\alpha,\beta)=(a_1\ \cdots\ a_n)\ G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}=\left[(a_1a_2\ \cdots\ a_n)G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}\right]'$$

$$=(b_1b_2\ \cdots b_n)G'\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}=(b_1b_2\ \cdots\ b_n)G\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}$$

$$=(\beta,\alpha).$$

2）$(k\alpha,\ \beta)=k(\alpha,\ \beta)$

事实上

$$(k\alpha,\ \beta)=(ka_1\ ka_2\ \cdots\ ka_n)G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

$$=k(a_1a_2\ \cdots\ a_n)\ G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}=k(\alpha,\beta).$$

3）$(\alpha+\beta,\ \gamma)=(a_1+b_1\ \cdots\ a_n+b_n)G\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}$

$$=[(a_1\ \cdots\ a_n)+(b_1\ \cdots\ b_n)]G\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}$$

$$=(a_1\ \cdots\ a_n)G\begin{pmatrix}c_1\\c_2\\\vdots\\c_n\end{pmatrix}$$

$$+(b_1\ \cdots\ b_n)G\begin{pmatrix}c_1\\\vdots\\c_n\end{pmatrix}=(\alpha,\ \gamma)+(\beta,\ \gamma)$$

其中，$\gamma=c_1\varepsilon_1+c_2\varepsilon_2+\cdots+c_n\varepsilon_n$

4）$(\xi,\xi)\geqslant 0$，只当 $\xi=0$ 时才有 $(\xi,\xi)=0$

事实上

$$(\xi,\xi)=(x_1,\ x_2\ \cdots\ x_n)\ G\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\xi=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n$$

因为G是正定矩阵，所以

$$(x_1\ x_2\ \cdots\ x_n)G\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\geqslant 0$$

只当 $x_1=0$，$x_2=0$，…，$x_n=0$ 时，才有

$$(x_1x_2\ \cdots\ x_n)\ G\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=0$$

这就是

$(\xi,\xi)\geqslant 0$，只当 $\xi=0$ 时才有 $(\xi,\xi)=0$.这样按定义，线性空间V用公式（2）做为内积运算公式构成了欧氏空间，而且基底 ε_1，ε_2，$\cdots$，ε_n 的度量矩阵恰是正定矩阵G。即

$$(\varepsilon_i,\ \varepsilon_j)=g_{ij}.$$

下面我们考虑任意两个基底的度量矩阵之间的关系。设ε_1，ε_2，$\cdots$，ε_n 与 δ_1，δ_2，$\cdots$，δ_n 是欧氏空间的两个基底，他们的度量矩阵分别为G与G_1。令过渡阵为P：

$$(\delta_1\ \cdots\ \delta_n)=(\varepsilon_1\ \cdots\ \varepsilon_n)P$$

若向量

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n=a_1'\delta_1+a_2'\delta_2+\cdots+a_n'\delta_n$$

$$\beta=b_1\varepsilon_1+b_2\varepsilon_2+\cdots+b_n\varepsilon_n=b_1'\delta_1+b_2'\delta_2+\cdots+b_n'\delta_n$$

则

$$P\begin{pmatrix}a_1'\\a_2'\\\vdots\\a_n'\end{pmatrix}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},\qquad P\begin{pmatrix}b_1'\\b_2'\\\vdots\\b_n'\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

即

$$(a_1\ a_2\ \cdots\ a_n)=(a_1'\ a_2'\ \cdots\ a_n')P'$$

$$(b_1\ b_2\ \cdots\ b_n)=(b_1'\ b_2'\ \cdots\ b_n')P'$$

于是

$$(\alpha,\beta)=(a_1a_2\ \cdots\ a_n)\ G\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}=(a_1'a_2'\ \cdots\ a_n')P'GP\begin{pmatrix}b_1'\\b_2'\\\vdots\\b_n'\end{pmatrix}$$

$$=(a_1'a_2'\ \cdots\ a_n')G_1\begin{pmatrix}b_1'\\b_2'\\\vdots\\b_n'\end{pmatrix}$$

得

$$G_1 = P'GP.$$

所以两个基底的度量矩阵是相合的．反之，如果G是欧氏空间V中基底 ε_1，ε_2，…，ε_n 的度量矩阵，那么对于凡与G相合的矩阵G_1

$$G_1 = P'GP$$

都有V的基底 δ_1，δ_2，…，δ_n 以 G_1 为度量矩阵．显然，用 P 为过渡阵，从 ε_1，ε_2，…，ε_n 得到的基底 δ_1，δ_2，…，δ_n 的度量矩阵就是G_1．这样我们就证明了

命题 1　设V是n维欧氏空间，那么 V 在各个基底上的度量矩阵恰为一类相合的正定矩阵．

例 1　设$V=[\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n]$是D上线性空间．$G_1$，$G_2$ 为任意两个不同的 n 阶正定矩阵．于是在V的基底 ε_1，ε_2，…，ε_n 上G_1与 G_2 各定义一个内积运算都使V成为欧氏空间，ε_1，ε_2，…，ε_n的度量矩阵分别为 G_1与G_2．因为 $G_1 \neq G_2$，所以二者定义的内积运算不全一样，这样，在同一基底 ε_1，…，ε_n 上以不同正定矩阵 G_1，G_2 为度量矩阵得到的是不同的欧氏空间．

另一方面，因为 G_2 与 G_1 必相合，即存在可逆矩阵P，使

$$G_2 = P'G_1P$$

于是以 P 为过渡阵得到V的另一基底

$$(\delta_1, \delta_2, \cdots, \delta_n) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)P$$

而在 δ_1，δ_2，…，δ_n 上的度量矩阵正是 G_2．这时，同一欧氏空间V，在两个不同基底 ε_1 …，ε_n 与 δ_1，…，δ_n 上的度量矩阵分别是G_1与G_2，这表明：任意两个不同的正定矩阵 G_1，G_2 都是同一个欧氏空间在某两个基底上的度量矩阵．

在有限维欧氏空间V里选取基底，对于该空间的线性运算来说有完全同等的意义，但是对于空间的内积运算来说，作用就大不相同了．由于内积的运算完全被一个基底的度量矩阵所确定，因此一个基底的度量矩阵G的形式越简单，那么内积运算在这个基底上的计算公式也就越简单．于是，就提出这样问题：在n 维欧氏空间V中怎样基底上的度量矩阵最简单？回想一下实对称阵在相合之下的

标准形．使知，在 n阶正定矩阵的相合类里，n 阶单位阵是该相合类中最简单的——标准形．于是我们假定 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是这样的一个基底，即它的度量矩阵是单位矩阵，这表明

$$(\varepsilon_i, \varepsilon_j)=\begin{cases}0, & i\neq j;\\ 1, & i=j.\end{cases}$$

从而基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是这样一组向量：任何不同两个向量都正交，每一个向量都是单位向量，即 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是一组标准正交向量．这使我们导至

定义 1 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是欧氏空间 V 的一个基底，如果 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是一组标准正交向量，则称 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是V的标准正交基底。

按定义自然有，$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是标准正交基底当且仅当$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 的度量矩阵是单位矩阵．这个事实的另一说法：$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是标准正交基底当且仅当在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 上的内积公式为

$$(\alpha, \beta)=a_1b_1+a_2b_2+\cdots+a_nb_n$$

其中 $\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$，$\beta=b_1\varepsilon_1+b_2\varepsilon_2+\cdots+b_n\varepsilon_n$.

命题 2 **有限维欧氏空间必有标准正交基底。**

证明 由于n 维欧氏空间一切基底的度量矩阵恰好是正定矩阵所构成的相合类．而这个相合类中含有单位阵在内，从而确有以单位阵为度量矩阵的基底存在，并且凡是以单位阵为度量矩阵的基底都是标准正交基底。

这个命题不仅肯定了标准正交基底的存在性，同时也提供了一个求出标准正交基底的方法。

例 2 在$D^{(2)}$中，设在基底 $\varepsilon_1=(1,0)$，$\varepsilon_2=(0,1)$ 上的度量矩阵为

$$G=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$$

求标准正交基底。

解：因为

$$\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

于是以

$$P=\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

为过渡阵从 ε_1，ε_2 得到另一基底

$$(\delta_1\ \ \delta_2)=(\varepsilon_1\ \ \varepsilon_2)P=(\varepsilon_1\ \varepsilon_2)\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

即

$$\delta_1=\varepsilon_1=(1,0)$$

$$\delta_2=\varepsilon_2-\varepsilon_1=(-1,1)$$

这个基底的度量矩阵是单位矩阵，所以δ_1，δ_2 是标准正交基底．比如，可以具体验证如下：

$$(\delta_1,\ \ \delta_1)=(1\ \ 0)\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}=(1\ \ 1)\begin{pmatrix} 1 \\ 0 \end{pmatrix}=1,$$

$$(\delta_1,\ \ \delta_2)=(1\ \ 0)\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix}=(1\ \ 1)\begin{pmatrix} -1 \\ 1 \end{pmatrix}=0,$$

$$(\delta_2,\ \ \delta_1)=(-1\ \ 1)\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}=(0\ \ 1)\begin{pmatrix} 1 \\ 0 \end{pmatrix}=0,$$

$$(\delta_2,\ \ \delta_2)=(-1\ \ 1)\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 1 \end{pmatrix}=(0\ \ 1)\begin{pmatrix} -1 \\ 1 \end{pmatrix}=1.$$

命题 3　设$V=[\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n]$ 是 n 维欧氏空间，α_1，α_2，$\cdots$，α_m 是一组标准正交向量，$m<n$，那么，存在 $\beta\in V$，使得 α_1，$\cdots$，α_m，β 也是标准正交组．

证　设 $\beta'=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n$，令

$$(\alpha_1,\ \beta')=0,\ \ (\alpha_2,\ \beta')=0,\ \ \cdots,\ \ (\alpha_m,\ \beta')=0$$

即

$$
\begin{cases}
k_1(\alpha_1,\varepsilon_1)+k_2(\alpha_1,\varepsilon_2)+\cdots+k_n(\alpha_1,\varepsilon_n)=0\\
k_1(\alpha_2,\varepsilon_1)+k_2(\alpha_2,\varepsilon_2)+\cdots+k_n(\alpha_2,\varepsilon_n)=0\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
k_1(\alpha_m,\varepsilon_1)+k_2(\alpha_m,\varepsilon_2)+\cdots+k_n(\alpha_m,\varepsilon_n)=0
\end{cases}
\tag{3}
$$

（3）是齐次线性方程组，系数阵为

$$
A=\begin{pmatrix}
(\alpha_1,\varepsilon_1) & (\alpha_1,\varepsilon_2) & \cdots & (\alpha_1,\varepsilon_n)\\
(\alpha_2,\varepsilon_1) & (\alpha_2,\varepsilon_2) & \cdots & (\alpha_2,\varepsilon_n)\\
\cdots & \cdots & \cdots & \cdots\\
(\alpha_m,\varepsilon_1) & (\alpha_m,\varepsilon_2) & \cdots & (\alpha_m,\varepsilon_n)
\end{pmatrix}
$$

由于 $m<n$，所以（3）有非零解，从而非零向量 β'，使 α_1，…，α_m，β' 是一个正交组，取

$$\beta=\frac{\beta'}{|\beta'|}$$

即得 α_1，α_2，…，α_m，β 是标准正交组。

推论 **若 α_1，α_2，…，$\alpha_m(m\leqslant n)$ 是V的任一标准正交组，则存在 β_{m+1}，…，β_n 使**

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_m,\ \beta_{m+1},\ \cdots,\ \beta_n$$

是V的标准正交基底。

证明　我们对 $n-m$ 作归纳法。

当 $n-m=0$ 时，α_1，α_2，…，α_m 就是标准正交基底，显然命题成立。

假设 $n-m=k$ 时成立。也就是说，当向量组 α_1，α_2，…，α_m 还差 k 个向量构成标准正交基底时，则存在k个向量 $\beta_1,\beta_2,\cdots,\beta_k$，使向量组 α_1，α_2，…，α_m，β_1，β_2，…，β_k 成为标准正交基底。

现证明 $n-m=k+1$ 时命题成立。显然 $m<n$，由命题3则有向量 β_1 使得 α_1，α_2，…，α_m,β_1 成为标准正交组，此时向量组α_1，α_2，…，α_m，β_1 还差 k个向量构成标准正交基底，由归纳假设知，存在k个向量 β_2，β_3，…，β_{k+1} 使向量组 α_1，α_2，…，α_m，β_1，β_2，

$\cdots$, β_{k+1} 成为V的标准正交基底，从而命题得证。

下面我们给出一种重要的求标准正交基底的方法。

命题 4 对欧氏空间V的任一基底 ε_1, ε_2, $\cdots$, ε_n 按下列方式能求得V的唯一的一个正交基底 δ_1, δ_2, $\cdots$, δ_n:

$$\begin{aligned}
&\delta_1=\varepsilon_1\\
&\delta_2=a_{12}\delta_1+\varepsilon_2\\
&\delta_3=a_{13}\delta_1+a_{23}\delta_2+\varepsilon_3\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&\delta_k=a_{1k}\delta_1+a_{2k}\delta_2+\cdots+a_{k-1\,k}\delta_{k-1}+\varepsilon_k\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
&\delta_n=a_{1n}\delta_1+a_{2n}\delta_2+\cdots+a_{n-1\,n}\delta_{n-1}+\varepsilon_n
\end{aligned}\tag{4}$$

从而把 δ_1, δ_2, $\cdots$, δ_n 标准化之后就得V的一个标准正交基底。

证 只须指出按着（4）的方式，依据正交性的要求恰能决定一组系数 a_{12}; a_{13}, a_{23}; $\cdots$; a_{1n}, $\cdots$, $a_{n-1\,n}$ 即可。

第一步，首先取定 $\delta_1=\varepsilon_1$。

第二步，令 $\delta_2=a_{12}\delta_1+\varepsilon_2$，依据 $\delta_1\perp\delta_2$，即用 $(\delta_1,\ \delta_2)=0$来决定系数 a_{12}。于是

$$\begin{aligned}
(\delta_1,\ \delta_2)&=(\delta_1,\ a_{12}\delta_1+\varepsilon_2)\\
&=(\delta_1,\ a_{12}\delta_1)+(\delta_1,\varepsilon_2)\\
&=a_{12}(\delta_1,\delta_1)+(\delta_1,\ \varepsilon_2)\\
&=0
\end{aligned}$$

从而即得

$$a_{12}=-\frac{(\delta_1,\varepsilon_2)}{(\delta_1,\delta_1)}.$$

第三步，令$\delta_3=a_{13}\delta_1+a_{23}\delta_2+\varepsilon_3$，依据$\delta_1\perp\delta_3$，$\delta_2\perp\delta_3$即用$(\delta_1,\ \delta_3)=0$，$(\delta_2,\ \delta_3)=0$ 来决定系数 a_{13}，a_{23}。于是

$$\begin{aligned}
(\delta_1,\ \delta_3)&=(\delta_1,\ a_{13}\delta_1+a_{23}\delta_2+\varepsilon_3)\\
&=(\delta_1,\ a_{13}\delta_1)+(\delta_1,\ a_{23}\delta_2)+(\delta_1,\ \varepsilon_3)\\
&=a_{13}(\delta_1,\ \delta_1)+a_{23}(\delta_1,\ \delta_2)+(\delta_1,\ \varepsilon_3)
\end{aligned}$$

$$= a_{13}(\delta_1, \delta_1) + (\delta_1, \varepsilon_3) = 0$$

$$\begin{aligned}(\delta_2, \delta_3) &= (\delta_2, a_{13}\delta_1 + a_{23}\delta_2 + \varepsilon_3)\\ &= a_{13}(\delta_2, \delta_1) + a_{23}(\delta_2, \delta_2) + (\delta_2, \varepsilon_3)\\ &= a_{23}(\delta_2, \delta_2) + (\delta_2, \varepsilon_3) = 0\end{aligned}$$

从而即得

$$a_{13} = -\frac{(\delta_1, \varepsilon_3)}{(\delta_1, \delta_1)}$$

$$a_{23} = -\frac{(\delta_2, \varepsilon_3)}{(\delta_2, \delta_2)}$$

这样继续作下去，到第 k 步，令

$$\delta_k = a_{1k}\delta_1 + a_{2k}\delta_2 + \cdots + a_{k-1k}\delta_{k-1} + \varepsilon_k$$

依据 $\delta_k \perp \delta_i$ $(i = 1, 2, \cdots, k-1)$ 便可唯一决定出

$$a_{1k} = -\frac{(\delta_1, \varepsilon_k)}{(\delta_1, \delta_1)}, \quad a_{2k} = -\frac{(\delta_2, \varepsilon_k)}{(\delta_2, \delta_2)}, \cdots, a_{k-1k} = -\frac{(\delta_{k-1}, \varepsilon_k)}{(\delta_{k-1}, \delta_{k-1})}$$

这样做到第 n 步就可求出满足关系式（4）的 V 的正交基底 δ_1, δ_2, $\cdots$, δ_n. 显然（4）中诸系数是由 ε_1, $\varepsilon_2, \cdots$, ε_n 唯一决定的。即按（4）的方式恰有一个正交基底存在．最后把 δ_1, δ_2, $\cdots$, δ_n 标准化，即令

$$\delta_i' = \frac{\delta_i}{|\delta_i|} \qquad i = 1, 2, \cdots, n$$

即得 V 的标准正交基底 δ_1', δ_2', $\cdots$, δ_n'.

不难看出，按方式（4）求出的正交基底有以下特性：

$$L(\delta_1, \delta_2, \cdots, \delta_k) = L(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k) \quad k = 1, 2, \cdots, n$$

例 3 已知欧氏空间 $D^{(3)}$，其基底 $\varepsilon_1 = (1, 0, 0)$ $\varepsilon_2 = (0, 1, 0)$, $\varepsilon_3 = (0, 0, 1)$ 的度量矩阵为

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

求 $D^{(3)}$ 的标准正交基底。

解　第一种方法．因为

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

所以以

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

为过渡阵得到的基底

$$(\delta_1 \delta_2 \delta_3) = (\varepsilon_1\ \varepsilon_2\ \varepsilon_3)\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$$

即

$$\delta_1 = \varepsilon_1$$

$$\delta_2 = \frac{1}{\sqrt{2}}\varepsilon_2$$

$$\delta_3 = \frac{1}{\sqrt{3}}\varepsilon_3$$

就是 $D^{(3)}$ 的一个标准正交基底．

下面用命题 4 的正交化方法，求出 $D^{(3)}$ 的一个标准正交基底

令　$\delta_1 = \varepsilon_1 = (1, 0, 0)$

$\delta_2 = a_{12}\delta_1 + \varepsilon_2 = (a_{12},\ 1,\ 0)$

$\delta_3 = a_{13}\delta_1 + a_{23}\delta_2 + \varepsilon_3 = (a_{13} + a_{23}a_{12},\ a_{23},\ 1)$

于是

$$a_{12} = -\frac{(\delta_1, \varepsilon_2)}{(\delta_1, \delta_1)} = 0,\ \text{其中}\ (\delta_1, \varepsilon_2) = (1\ \ 0\ \ 0)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 0,$$

所以 $\delta_2=\varepsilon_2$，

$$a_{13}=-\frac{(\delta_1,\varepsilon_3)}{(\delta_1,\delta_1)}=0，\text{其中}(\delta_1,\varepsilon_3)=(1\ \ 0\ \ 0)\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}=0,$$

$$a_{23}=-\frac{(\delta_2,\varepsilon_3)}{(\delta_2,\delta_2)}=0，\text{其中}(\delta_2,\varepsilon_3)=(0\ \ 1\ \ 0)\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}=0.$$

所以 $\delta_3=\varepsilon_3$。

这样，由 ε_1，ε_2，ε_3 得到的正交基底就是原来的 ε_1，ε_2，ε_3(因为 ε_1，ε_2，ε_3 原来已经是正交的）。再把这个正交基底标准化之后就得到了标准正交基底：

$$\delta_1'=\frac{\delta_1}{|\delta_1|},\quad \delta_2'=\frac{\delta_2}{|\delta_2|},\quad \delta_3'=\frac{\delta_3}{|\delta_3|}.$$

命题 5　设 ε_1，ε_2，$\cdots$，ε_n 是V的标准正交基底，

$$(\delta_1,\ \delta_2,\ \cdots,\ \delta_n)=(\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n)T,\quad T=(t_{ij})$$

那么 δ_1，δ_2，$\cdots$，δ_n 是标准正交基底当且仅当T是正交阵。

证　因为 ε_1，ε_2，$\cdots$，ε_n 是标准正交基底，即有

$$(\varepsilon_i,\ \varepsilon_j)=\begin{cases}0, & i\neq j;\\1, & i=j.\end{cases}$$

于是

$$(\delta_i,\ \delta_j)=t_{1i}t_{1j}+t_{2i}t_{2j}+\cdots+t_{ni}t_{nj}$$

从而

$$(\delta_i,\ \delta_j)=\begin{cases}0, & i\neq j;\\1, & i=j,\end{cases}$$

当且仅当

$$t_{1i}t_{1j}+t_{2i}t_{2j}+\cdots+t_{ni}t_{nj}=\begin{cases}0, & i\neq j;\\1, & i=j.\end{cases}$$

这说明 δ_1，δ_2，$\cdots$，δ_n 是标准正交基底当且仅当 $T=(t_{ij})$ 是正交

阵.

设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是V的任一标准正交基底，对V中任意向量

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n$$

$$\beta = b_1\varepsilon_1 + b_2\varepsilon_2 + \cdots + b_n\varepsilon_n$$

则有

$$(\alpha, \beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

由此可见，对任意的 n 维欧氏空间来说，从它的标准正交基底来看，不仅它们的线性运算是完全相同的，即都是坐标的线性运算，而且内积运算也是一致的，即都是坐标乘积之和。为了明确这种现象的实质，我们给出以下的定义.

定义 2　设 V_1, V_2 都是欧氏空间，φ 是作为线性空间V_1与V_2的同构映射，如果还具有性质：若

$$\varphi: \alpha \longmapsto \alpha', \quad \beta \longmapsto \beta'$$

则

$$(\alpha, \beta) = (\alpha', \beta') \qquad \forall \alpha, \beta \in V_1$$

那么则说 φ 是欧氏空间 V_1 与V_2 的同构映射，这时称V_1 与V_2是同构的，还是记作 $V_1 \simeq V_2$.

显然单满映射φ是V_1 与V_2的同构映射，当且仅当φ具有性质：

1) $\varphi(\alpha + \beta) = \varphi(\alpha) + \varphi(\beta)$

2) $\varphi(k\alpha) = k\varphi(\alpha)$

3) $(\varphi(\alpha), \varphi(\beta)) = (\alpha, \beta)$.

命题 6　设V_1 与 V_2 都是有限维的欧氏空间，那么$V_1 \simeq V_2$当且仅当 $\dim V_1 = \dim V_2$.

证. 必要性是显然的. 因为两个欧氏空间 V_1 与 V_2 同构，首先做为线性空间二者是同构的，所以二者的维数相等. 下而证明充分性.

设 $\dim V_1 = \dim V_2$，于是可在 V_1 与 V_2 中分别取标准正交基底

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$$

$$\varepsilon_1', \ \varepsilon_2', \ \cdots, \ \varepsilon_n'$$

规定V_1到V_2的映射

$$\varphi: \ \xi \longmapsto \xi'$$

其中 $\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n$, $\xi' = x_1\varepsilon_1' + x_2\varepsilon_2' + \cdots + x_n\varepsilon_n'$ 显然φ是V_1 到 V_2 的一个映射，而且是单的、满的，即φ是 V_1 与 V_2 的一一对应。而且具有性质：如果

$$\varphi: \ \alpha \longmapsto \alpha', \qquad \beta \longmapsto \beta'$$

于是令

$$\alpha = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n, \ \ \beta = b_1\varepsilon_1 + b_2\varepsilon_2 + \cdots + b_n\varepsilon_n$$

则有

$$\alpha' = a_1\varepsilon_1' + a_2\varepsilon_2' + \cdots + a_n\varepsilon_n', \beta' = b_1\varepsilon_1' + b_2\varepsilon_2' + \cdots + b_n\varepsilon_n'$$

从而

1) $$\begin{aligned}\varphi(\alpha+\beta) &= \varphi((a_1+b_1)\varepsilon_1 + (a_2+b_2)\varepsilon_2 + \cdots + (a_n+b_n)\varepsilon_n) \\ &= (a_1+b_1)\varepsilon_1' + (a_2+b_2)\varepsilon_2' + \cdots + (a_n+b_n)\varepsilon_n' \\ &= a_1\varepsilon_1' + a_2\varepsilon_2' + \cdots + a_n\varepsilon_n' + b_1\varepsilon_1' + b_2\varepsilon_2' + \cdots + b_n\varepsilon_n' \\ &= \alpha' + \beta'\end{aligned}$$

即

$$\varphi: \ \alpha+\beta \longmapsto \alpha' + \beta'$$

2) $$\begin{aligned}\varphi(k\alpha) &= \varphi[(ka_1)\varepsilon_1 + (ka_2)\varepsilon_2 + \cdots + (ka_n)\varepsilon_n] \\ &= (ka_1)\varepsilon_1' + (ka_2)\varepsilon_2' + \cdots + (ka_n)\varepsilon_n' \\ &= k(a_1\varepsilon_1' + a_2\varepsilon_2' + \cdots + a_n\varepsilon_n') \\ &= k\alpha'\end{aligned}$$

即

$$\varphi: \ k\alpha \longmapsto k\alpha'$$

3) $(\alpha, \beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n = (\alpha', \beta')$.

所以φ是 V_1 与V_2 的同构映射，即 $V_1 \simeq V_2$。

命题6 表明，对有限维欧氏空间的代数结构来说决定性的条件仅仅是欧氏空间的维数。从而，只有维数不同的欧氏空间才是不同的，而维数相同的欧氏空间，没有实质性的差别。

练 习 二

1．在欧氏空间$D^{(3)}$中，按通常的数积定义，设 $\alpha_1=(1,0,0)$，$\alpha_2=(1,1,0)$，$\alpha_3=(1,1,1)$，求

1）由α_1，α_2，α_3 所构成的基底的度量矩阵，

2）用两种方法求$D^{(3)}$ 的标准正交基底。

2．在欧氏空间$D^{(4)}$ 中，已知基底

$\alpha_1=(1,-1,0,0)$，$\alpha_2=(-1,2,0,0)$，$\alpha_3=(0,1,2,1)$，$\alpha_4=(1,0,1,1)$ 的度量矩阵为

$$B=\begin{pmatrix} 2 & -3 & 0 & 1 \\ -3 & 6 & 0 & -1 \\ 0 & 0 & 13 & 9 \\ 1 & -1 & 9 & 7 \end{pmatrix}$$

求基底

$\varepsilon_1=(1,0,0,0)$，$\varepsilon_2=(0,1,0,0)$，$\varepsilon_3=(0,0,1,0)$，$\varepsilon_4=(0,0,0,1)$ 的度量矩阵。

3．在欧氏空间V中，基底 ε_1，ε_2，…，ε_n 是标准正交的必要而且只要对V中任一向量

$$\alpha=a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n$$

总有

$$(\alpha,\ \varepsilon_i)=a_i \qquad i=1,2,\cdots,n.$$

4．求齐次线性方程组

$$2x_1+x_2-x_3+x_4-3x_5=0$$

$$x_1+x_2-x_3+x_5=0$$

的解空间（作为 $D^{(5)}$的子空间）的标准正交基底。

5．设 α_1，α_2，…，α_m 是欧氏空间 V 的一个标准正交组，证明，对任意 $\alpha\in V$，以下不等式成立：

$$\sum_{i=1}^{m}(\alpha,\alpha_i)^2\leqslant|\alpha|^2$$

6. 设 ε_1, ε_2, ε_3 是欧氏空间V的标准正交基底，证明：

$$\alpha_1=\frac{1}{3}(2\varepsilon_1+2\varepsilon_2-\varepsilon_3)$$

$$\alpha_2=\frac{1}{3}(2\varepsilon_1-\varepsilon_2+2\varepsilon_3)$$

$$\alpha_3=\frac{1}{3}(\varepsilon_1-2\varepsilon_2-2\varepsilon_3)$$

仍是V的标准正交基底。

7. 设 ε_1, ε_2, ε_3, ε_4, ε_5 是欧氏空间V的一组标准正交基底，令

$$V_1=L(\alpha_1,\ \alpha_2,\ \alpha_3)$$

其中

$$\alpha_1=\varepsilon_1+\varepsilon_5,\quad \alpha_2=\varepsilon_1-\varepsilon_2+\varepsilon_4,\quad \alpha_3=2\varepsilon_1+\varepsilon_2+\varepsilon_3$$

求 V_1 的一组标准正交基底。

§3 对称变换与正交变换

本节讨论有限维欧氏空间的两种线性变换，它们都有明显的几何意义，同时与空间的内积运算有密切的联系。

设V是 n维欧氏空间，ε_1, ε_2, …, ε_n 是 V 的一个标准正交基底。考虑V的一个线性变换 σ。

如果 σ 在标准正交基底 ε_1, ε_2, …, ε_n 上的表示矩阵A是对称矩阵，则称 σ是对称变换。

如果 σ 在标准正交基底 ε_1, ε_2, …, ε_n 上的表示矩阵A是正交矩阵，则称 σ是正交变换。

例 1 $D^{(3)}$ 为按通常的内积运算构成的欧氏空间，$\varepsilon_1=(1,\ 0,$

0)，$\varepsilon_2=(0,1,0)$，$\varepsilon_3=(0,0,1)$ 是一个标准正交基底，σ_1，σ_2，σ_3 是 $D^{(3)}$ 的三个线性变换，它们在 ε_1，ε_2，ε_3 上的表示矩阵分别为

$$A_1=\begin{pmatrix}-1 & & \\ & 1 & \\ & & 1\end{pmatrix}\quad A_2=\begin{pmatrix}1 & & \\ & 0 & -1 \\ & 1 & 0\end{pmatrix}\quad A_3=\begin{pmatrix}2 & & \\ & 2 & \\ & & 2\end{pmatrix}$$

于是，因为A_1 是对称矩阵，所以 σ_1 是对称变换。A_2 是正交矩阵，所以 σ_2 是正交变换。A_3 是对称阵，所以 σ_3 是对称变换。显然 A_1 也是正交阵，所以 σ_1 也是正交变换。容易看出这三个线性变换的几何意义，σ_1 是关于 yz——坐标面的对称，σ_2 是绕x—轴的旋转，旋转角为$\frac{\pi}{2}$。σ_3 是以 2 为相似系数的相似变换。

命题 1　n 维欧氏空间V的线性变换σ是对称变换当且仅当σ在任一标准正交基底上的表示矩阵都是对称阵。

证　充分性是自然的，下面证明必要性。

因为σ 是对称变换，可取某一标准正交基底 ε_1，ε_2，…，ε_n 使 σ 的表示矩阵 A 是对称矩阵，令 δ_1，δ_2，…，δ_n 是任一标准正交基底，由 ε_1，ε_2，…，ε_n 到 δ_1，δ_2，…，δ_n 的过渡阵为正交阵P，即

$$(\delta_1,\ \delta_2,\ \cdots,\ \delta_n)=(\varepsilon_1,\ \varepsilon_2,\ \cdots,\ \varepsilon_n)P$$

于是 σ 在 δ_1，δ_2，…，δ_n 上的表示矩阵

$$B=P^{-1}AP$$

从而

$$B'=(P^{-1}AP)'=P'A'(P^{-1})'=P^{-1}AP=B$$

故B是对称阵，命题 1 得证。

命题 2　n 维欧氏空间V的线性变换 σ 是对称变换当且仅当 $(\sigma(\alpha),\beta)=(\alpha,\sigma(\beta))$，$\forall\alpha,\beta\in V$。

证　任取V 的标准正交基底 ε_1，ε_2，…，ε_n，σ 在此基底上的表示矩阵为 $A=(a_{ij})$。

先证必要性．用$\overline{\alpha}$，$\overline{\beta}$表示向量α，β在标准正交基底ε_1，ε_2，…，ε_n上的坐标列，于是

$$(\sigma(\alpha),\beta)=(\overline{\sigma(\alpha)})'\cdot\overline{\beta}=(A\overline{\alpha})'\overline{\beta}=\overline{\alpha}'A'\overline{\beta}=\overline{\alpha}'A\overline{\beta}$$

$$=\overline{\alpha}'\overline{\sigma(\beta)}=(\alpha,\sigma(\beta))。$$

再证充分性．

$$a_{ij}=(\sigma(\varepsilon_j),\varepsilon_i)=(\varepsilon_j,\sigma(\varepsilon_i))=a_{ji}。$$

即$A'=A$，所以σ在标准正交基底ε_1，ε_2，…，ε_n上的表示矩阵是对称的，故σ是对称变换．

定理1　设σ为对称变换，那么存在标准正交基底使σ的表示矩阵是对角阵．

证　任取标准正交基底ε_1，ε_2，…，ε_n　σ的表示矩阵为A，由第十章§5定理2知任一实对称阵必正交相似于对角阵$\text{diag}\{a_1, a_2, \cdots, a_n\}$．$a_1$，$a_2$，…，$a_n$　恰是A的全部特征根．这样，存在正交阵T，使

$$T^{-1}AT=\text{diag}\{a_1,\ a_2,\ \cdots,a_n\}$$

于是以T为过渡阵，从标准正交基底ε_1，ε_2，…，ε_n得另一标准正交基底δ_1，δ_2，…，δ_n：

$$(\delta_1\delta_2\ \cdots\ \delta_n)=(\varepsilon_1\varepsilon_2\ \cdots\ \varepsilon_n)T$$

而σ在δ_1，δ_2，…，δ_n上的表示矩阵就是对角阵$\text{diag}\{a_1, a_2, \cdots, a_n\}$．

推论　n维欧氏空间V的线性变换σ是对称的当且仅当σ有n个两两正交的一维不变子空间W_1，W_2，…，W_n使

$$V=W_1\oplus W_2\oplus\cdots\oplus W_n。$$

命题3　n维欧氏空间的线性变换σ是正交变换当且仅当σ在任一标准正交基底上的表示矩阵都是正交阵。

证　充分性是自然的，下面证明必要性。

因σ是正交变换，可取某一标准正交基底$\{\varepsilon_i\}$使σ的表示矩阵

A是正交阵，令 $\delta_1, \delta_2, \cdots, \delta_n$ 是V的任一标准正交基底，由 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 到 $\delta_1, \delta_2, \cdots, \delta_n$ 的过渡阵是P，即

$$(\delta_1 \delta_2 \cdots \delta_n) = (\varepsilon_1 \ \varepsilon_2 \ \cdots \ \varepsilon_n) P$$

于是σ在 $\delta_1, \delta_2, \cdots, \delta_n$ 上的表示矩阵

$$B = P^{-1} A P$$

因为 P, A, P^{-1} 都是正交阵，所以 B 也是正交阵，从而命题得证.

命题 4 n **维欧氏空间** V **的线性变换** σ **是正交变换当且仅当** $(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta) \quad \forall \alpha, \beta \in V$.

证 任取 V 的标准正交基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$, σ 的表示矩阵为 $A = (a_{ij})$.

先证必要性. 用$\overline{\alpha}$, $\overline{\beta}$表示向量 α, β 在标准正交基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 上的坐标列. 于是

$$(\sigma(\alpha), \sigma(\beta)) = \overline{\sigma(\alpha)}' \cdot \overline{\sigma(\beta)} = (A\overline{\alpha})' A \overline{\beta} = \overline{\alpha}' A' A \overline{\beta}$$
$$= \overline{\alpha}' \overline{\beta} = (\alpha, \beta).$$

再证充分性，由

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)$$

则有

$$(\sigma(\varepsilon_i), \sigma(\varepsilon_j)) = (\varepsilon_i, \varepsilon_j) = \begin{cases} 0, & \text{当 } i \neq j \text{ 时}; \\ 1, & \text{当 } i = j \text{ 时}. \end{cases}$$

这说明 $\sigma(\varepsilon_1), \sigma(\varepsilon_2), \cdots, \sigma(\varepsilon_n)$ 也是V的标准正交基底，而

$$(\sigma(\varepsilon_1) \ \sigma(\varepsilon_2) \ \cdots \ \sigma(\varepsilon_n)) = (\varepsilon_1 \ \varepsilon_2 \ \cdots \ \varepsilon_n) A$$

由此可知A是正交矩阵，即 σ是正交变换. 命题得证.

例 2 设 $(\sigma(\alpha), \sigma(\alpha)) = (\alpha, \alpha)$，证明 σ是正交变换.

证 由命题 4 只要证明对任意 α, β 都有 $(\sigma(\alpha), \sigma(\beta))$ 即可. 为此分别计算

$$(\sigma(\alpha+\beta), \sigma(\alpha+\beta)) = (\sigma(\alpha)+\sigma(\beta), \sigma(\alpha)+\sigma(\beta))$$
$$= (\sigma(\alpha), \sigma(\alpha)) + 2(\sigma(\alpha), \sigma(\beta)) + (\sigma(\beta), \sigma(\beta))$$

$$= (\alpha, \alpha) + 2(\sigma(\alpha), \sigma(\beta)) + (\beta, \beta)$$

$$(\alpha+\beta, \alpha+\beta) = (\alpha, \alpha) + 2(\alpha, \beta) + (\beta, \beta)$$

而

$$(\sigma(\alpha+\beta), \sigma(\alpha+\beta)) = (\alpha+\beta, \alpha+\beta)$$

于是

$$2(\sigma(\alpha), \sigma(\beta)) = 2(\alpha, \beta)$$

从而便得

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta).$$

即 σ 是正交变换．此例说明：欧氏空间的线性变换 σ 如果不改变每一向量的长，那么 σ 就是正交变换，当然反过来自然也是对的。

定理 2　设 σ 为正交变换，那么存在标准正交基底，使 σ 的表示矩阵为

$$B = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_r & & & \\ & & & A_1 & & \\ & & & & \ddots & \\ & & & & & A_s \end{pmatrix} \tag{1}$$

其中 $a_i = \pm 1$ $(i=1, \cdots, r)$，A_j $(j=1, \cdots, s)$ 为二阶正交阵，且没有实特征根。

证　任取标准正交基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，$\sigma$ 的表示矩阵为 A，由第十章§6定理可知任一正交阵必正交相似于形式如（1）的阵，这样存在正交阵 T 使

$$B = T^{-1}AT$$

于是以 T 为过渡阵从标准正交基底 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 得另一个标准正交基底 $\delta_1, \delta_2, \cdots, \delta_n$：

$$(\delta_1\ \delta_2\ \cdots\ \delta_n) = (\varepsilon_1\ \varepsilon_2\ \cdots\ \varepsilon_n)T$$

而 σ 在 $\delta_1, \delta_2, \cdots, \delta_n$ 上的表示矩阵就是 B．从而定理得证。

推论　n 维欧氏空间 V 的线性变换 σ 是正交变换当且仅当 σ 有一组维数不大于 2 的两两正交的不变子空间 $W_1, W_2, \cdots, W_r$，

$V_1, V_2, \cdots, V_s$ 使得

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_r \oplus V_1 \oplus V_2 \oplus \cdots \oplus V_s$$

而 σ 在 $W_i (i = 1, 2, \cdots, r)$ 上的导出变换是相似系数 $a_i = \pm 1$ 的相似变换，σ 在 $V_j (j = 1, 2, \cdots, s)$ 上的导出变换是没有特征向量的旋转变换。

练 习 三

1. 证明，两个正交变换的乘积仍是正交变换。一个正交变换的逆变换还是正交变换。

2. 设 η 是欧氏空间 V 的一个单位向量，定义

$$\sigma(\alpha) = \alpha - 2(\eta, \alpha)\eta$$

证明：1）σ 是正交变换，这样的变换称为镜面反射。

2）σ 在标准正交基底的表示矩阵的行列式值是 -1。

3）在 n 维欧氏空间 V 中，如果正交变换 σ 的特征根为 1 的特征子空间的维数是 $n-1$ 维的，那么 σ 是镜面反射。

3. 设 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$ 与 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 是欧氏空间 V 的两个标准正交组，证明存在一个正交变换 σ 使得

$$\sigma(\varepsilon_i) = \alpha_i \qquad i = 1, 2, \cdots, m.$$

4. 设 α, β 是欧氏空间 V 中长度相同的非零向量，证明必有正交变换 σ 使得

$$\sigma(\alpha) = \beta.$$

5. 设 σ 是欧氏空间 V 的一个变换，如果对 V 中任意向量 α, β 都有

$$(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)$$

证明 σ 是正交变换。

6. 设 W 是欧氏空间 V 上的正交变换 σ 的不变子空间，证明 $W^{\perp}$ 也是 σ 的不变子空间。

7．设 σ_1，σ_2 是欧氏空间V的对称变换，那么 $\sigma_1\sigma_2+\sigma_2\sigma_1$也是$V$的对称变换。

8．设W是欧氏空间V的对称变换 σ的不变子空间，证明$W^\perp$是 σ的不变子空间。

习 题 十三

1．在 $D_4[x]$ 中，定义内积为

$$(f(x),\ g(x))=\int_{-1}^{1} f(x)g(x)dx$$

求出基底 1，x，x^2，x^3 而得到的一组标准正交基底。

2．设 α，β 是欧氏空间V中两个线性无关的向量，如果

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \text{ 与 } \frac{2(\alpha,\beta)}{(\beta,\beta)}$$

都是不大于 0 的整数，证明 α，β 的夹角只能是 $\frac{\pi}{2}$，$\frac{2}{3}\pi$，$\frac{3}{4}\pi$，$\frac{5}{6}\pi$。

3．设 W_1，W_2 是欧氏空间V的子空间，证明

1）$(W_1+W_2)^\perp=W_1^\perp\cap W_2^\perp$

2）$(W_1\cap W_2)^\perp=W_1^\perp+W_2^\perp$

4．设 α_1，α_2，$\cdots$，α_n 是欧氏空间V的 n个向量，

$$G(\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n)=\begin{vmatrix} (\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & \cdots & (\alpha_1,\alpha_n) \\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & \cdots & (\alpha_2,\alpha_n) \\ \cdots & \cdots & \cdots & \cdots \\ (\alpha_n,\alpha_1) & (\alpha_n,\alpha_2) & \cdots & (\alpha_n,\alpha_n) \end{vmatrix}$$

证明：α_1，α_2，$\cdots$，α_n 是线性无关的当且仅当 $G(\alpha_1,\ \alpha_2,\cdots,\ \alpha_n)\neq 0$。其中行列式 $G(\alpha_1,\alpha_2,\cdots,\ \alpha_n)$ 叫做 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的克拉姆行列式。

5．设 α_1，α_2，$\cdots$，α_n 是欧氏空间 V 的 n 个向量，ε_1，ε_2，

…，ε_n 是V的标准正交基，若

$$\alpha_i = a_{i1}\varepsilon_1 + a_{i2}\varepsilon_2 + \cdots + a_{in}\varepsilon_n \quad (i = 1, 2, \cdots, n)$$

则

$$G(\alpha_1, \alpha_2, \cdots, \alpha_n) = |a_{ij}|^2.$$

6．设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是欧氏空间V中的一个基底，$\beta_1, \beta_2, \cdots, \beta_n$ 是由 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 正交化而得到的一个正交基底．证明

$$G(\alpha_1, \alpha_2, \cdots, \alpha_n) = G(\beta_1, \beta_2, \cdots, \beta_n)$$
$$= (\beta_1, \beta_1)(\beta_2, \beta_2) \cdots (\beta_n, \beta_n).$$

7．若集合 S_1，S_2 正交，则 $S_1 \cap S_2$ 或是空集或是仅有一个零向量．

8．设

$$W_1 = L(\alpha_1, \alpha_2, \cdots, \alpha_m)$$
$$W_2 = L(\beta_1, \beta_2, \cdots, \beta_n)$$

如果

$$(\alpha_i, \beta_j) = 0 \quad i = 1, 2, \cdots m; \ j = 1, 2, \cdots, n.$$

那么 W_1，W_2 正交．

9．设V是欧氏空间，W是V的子空间，由分解式

$$V = W \oplus W^{\perp}$$

知V中任意向量 α都可唯一分解成

$$\alpha = \alpha_1 + \alpha_2$$

其中 $\alpha_1 \in W$，称 α_1 为 α在W上的内射影，$\alpha_2 \in W^{\perp}$ 称 α_2 为 α 在W上的外射影．设 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 是W的一个标准正交基底，证明 α 在W上的内射影为

$$(\alpha, \alpha_1)\alpha_1 + (\alpha, \alpha_2)\alpha_2 + \cdots + (\alpha, \alpha_m)\alpha_m.$$

10．证明：向量 $\beta \in W$ 是欧氏空间V中向量 α 在子空间W的内射影的充分必要条件是对任意 $\xi \in W$，

$$|\alpha - \beta| \leqslant |\alpha - \xi|.$$

11．设A是正交变换 σ在标准正交基底上的表示矩阵，且$|A| = -1$，则 σ一定以-1为特征根．

12. 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 和 $\beta_1, \beta_2, \cdots, \beta_n$ 是欧氏空间V的两组标准正交基底，如果V的一个正交变换 τ使得

$$\tau(\alpha_1)=\beta_1$$

那么

$$L(\tau(\alpha_2), \tau(\alpha_3), \cdots, \tau(\alpha_n))=L(\beta_2, \beta_3, \cdots, \beta_n).$$

13. 设 σ是欧氏空间V的一个线性变换，如果在某一个标准正交基底上的表示矩阵是反对称矩阵，则称 σ是反对称变换. 证明：

1）σ 是反对称变换的充分必要条件是

$$(\sigma(\alpha), \beta)=-(\alpha, \sigma(\beta))$$

2）W_1 是反对称变换 σ的不变子空间，则 $W_1^{\perp}$ 也是 σ 的不变子空间。

14. 1）设 α, β 是欧氏空间V的两个不同的单位向量，证明存在一个镜面反射 σ，使得

$$\sigma(\alpha)=\beta$$

2）证明：n 维欧氏空间V中任一正交变换 σ都可表成一系列镜面反射的乘积。

15. 设 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 和 $\beta_1, \beta_2, \cdots, \beta_m$ 是 n维欧氏空间V的两组向量，证明存在一个正交变换 σ使得

$$\sigma(\alpha_i)=\beta_i$$

的充分必要条件是

$$(\alpha_i, \alpha_j)=(\beta_i, \beta_j).$$

16. 设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

是实 n阶方阵，证明线性方程组

$$A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix} \tag{1}$$

有解的充分必要条件是向量 $\beta=(b_1, b_2, \cdots, b_n)\in D^{(n)}$与齐次线性方程组

$$A'\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix} \tag{2}$$

的解空间正交。

学 习 指 导

第七章　集合与映射

〔内容提要〕

结合前几章的内容，本章首先阐述了高等代数的一些特点，它们主要体现了高等代数讨论问题的方法论方面的特征。

本章讲三个主要问题：映射、代数运算、集的分解．而映射是基本的核心概念，对理解全章内容有决定性的作用．

分类的概念对高等代数来说，应该放在突出的重要位置上，尤其是在以下几章的讨论中，是不可缺少的．它既是解决问题的方法，又体现了讨论问题的一种基本观念．

整章是以下各章的基础，篇幅不多概念不少，而且大部分基本概念都有概括性、灵活性和抽象性等特点．因此，对于理解、把握这些概念并能初步运用它们是有一定难度的．这就要求读者予以适当的注意，尽可能把本章学习得具体些、灵活些，为以下几章的学习做好必要的准备．熟悉掌握一批恰当的例题是学好本章的非常重要的手段．

本章的主要概念是：

积集；

映射、单映射、满映射、单满映射（一一对应）、逆映射；

变换、单变换、满变换、单满变换（一一变换）、逆变换．

代数运算；

集合的子集分解（分类）、关系、等价关系．

本章的主要结果是：

关于结合律、交换律、分配律的三个命题；

分类与等价关系之间的联系（§4定理）、分类与映射之间的联系（§4命题）。

本章的主要方法是：

1．在两个集合之间给出映射；

2．建立代数运算；

3．给集合做成分类；

4．在集合上建立关系、等价关系。

通过这一章的学习，要求读者：

1．领会高等代数讨论问题的主要特点；

2．记住并集、交集、差集、补集、积集、幂集这些概念；

3．理解映射的一般概念，掌握一些基本例子；

4．理解代数运算的一般概念，掌握一些基本例子；

5．理解分类的一般概念，掌握一些基本例子；

6．理解等价关系的一般概念，掌握一些基本例子；

7．掌握分类与等价关系之间的联系，并能用来说明某些问题；

8．能够运用映射、分类、等价关系的一般概念分析、解释一定的问题。

〔内容分析〕

§1 集　合

本节内容比较简单，主要是在已知集合、子集、空集等概念的基础上，进而介绍了并集、交集、差集、补集、积集、幂集这些概念。所有这些概念都是容易掌握的，其中并集和交集比较常用。积集的概念是相当重要的，也常用到。特别是有一点须要明确的，即在积集$A\times A$中，当 $a\neq b$ 时，(a,b) 与 (b,a) 是不同的两个元素，

其中 $a,b\in A$. 诚然，这对我们是早有所知的.

关于幂集，我们已经知道，对于含有 n个元素的集合A，其幂集 $P(A)$ 恰好含有 2^n 个元素，即A共有 2^n 个子集. 换句话说就是，含 n个元素的集合A的幂集 $P(A)$ 的元素个数为 2 的 n次幂：2^n. 这似乎是称 $P(A)$ 为A的幂集的一种背景.

我们只规定了有限个集合的并集、交集的概念. 完全类似地，可以定义任意一些集合的并集和交集.

设 $W=\{A_1, A_2, \cdots\}$，其中$A_1, A_2, \cdots$ 都是任意的集合，于是可以确定以下两个集合：

$$S=\{x| \text{有某一 } A_i\in W, \text{使 } x\in A_i\},$$

$$T=\{x| \forall A_i\in W, x\in A_i\}.$$

称集合S为 $A_1, A_2, \cdots$ 的并集，T 为 $A_1, A_2, \cdots$ 的交集. 分别记作

$$S=\bigcup_{A_i\in W} A_i, \quad T=\bigcap_{A_i\in W} A_i.$$

例如 $D^{(2)}=\{(x,y)| x,y\in D\}$ 可以看做平面上一切点的集合. 考虑 $D^{(2)}$ 如下的子集：

$$l_a=\{(x,a)| x\in D\}, \quad \forall a\in D.$$

这个 l_a 就是平面上的一条水平直线. 于是

$$\bigcup_{a\in D} l_a=D^{(2)}, \quad \bigcap_{a\in D} l_a=\phi.$$

再如，考虑一元多项式的集合 $F[x]$，及其如下的子集：

$$d_i=\{f(x)\in F[x]| f(x)=a_0+a_1x+\cdots+a_{i-1}x^{i-1}\neq 0\}.$$

这里，d_i 就是一切次数小于 i 的多项式组成的子集. 于是

$$\bigcup_{i\in N} d_i=F[x]-\{0\}, \quad \bigcap_{i\in N} d_i=d_1.$$

如果考虑差集 $d_{i+1}-d_i$ 即次数等于 i 的一切多项式的子集，则有

$$\bigcup_{i\in N}(d_{i+1}-d_i)=F[x]-F, \quad \bigcap_{i\in N}(d_{i+1}-d_i)=\phi.$$

§2 映　射

映射概念的重要性已一再强调过了，这里不再重复。

学习本节（以下两节也如此）应着重解决两方面的问题。一方面是理解好概念，另一方面是掌握好例子，而且二者是互为因果的。

关于概念，首先是注意准确性。由于每一概念都有自身的特殊规定性，因此我们对于每一概念的内容必须有十分准确的认识，不可似是而非，更不容许有丝毫的误解。其次，注意把不同的概念之间的差异区分清楚。本节的概念较多，有些概念的含义又相当接近。这样就要求读者把那些含义相近的概念之间的异同明确起来，划清界限以免混同。

掌握一批恰当的例子是学好本章，不仅是本节，的必要手段。因为一方面必须把那些相当概括而又比较抽象的概念落实为具体的例子，另一方面也须要通过适当的例子来说明一些概念的内容及值得注意的地方。同时也是为了检验我们对于概念的理解是否准确、完整、透彻、灵活。

变换——集合A到自身的映射，是个非常重要的概念。在以下几章里都要用到它，而且专有一章集中讨论变换的问题。

在内容提要部分，我们提出过这样的要求：理解映射（变换）的一般概念，并能对给定的集合A与B建立它们之间的某些映射。解决好这个问题，关键在于我们对已给的集合A与B的了解如何。其实，在A与B之间建立这样或那样的映射，不外都是反映我们对A与B之间的关系的一些认识。完全可以这样说，对A与B之间的关系有怎样的认识就能建立起A与B之间的怎样的映射。例如，已知$M_2(D)$与D这样两个集合。所谓建立从$M_2(D)$到D的映射，其实就是使二阶方阵与实数之间建立起某种联系。因此，一般来说这完全取决于我们对于二阶方阵与实数之间的关系有怎样的认识。如果我们一点不了解二阶方阵与实数之间有何联系，那么我们根本

没有办法给出 $M_2(D)$ 与D的映射．诚然，我们都已知道，二阶方阵与数不仅有联系，而且有许多密切的有意义的重要联系．如每个二阶方阵都有行列式，其值就是D中的一个确定的数．再如，每个二阶方阵都有秩，这也是D中的一个确定的数．这样，我们相应的有两种反映这种联系的映射：

$$\varphi_1:\ (a_{ij}) \longmapsto \det(a_{ij}),$$

$$\varphi_2:\ (a_{ij}) \longmapsto \mathrm{rank}(a_{ij}).$$

由于目前我们还没认识到二阶方阵与实数之间另外的比较有意义的联系，因此，也就不能再给出 $M_2(D)$ 与D的比较有意义的映射．不过退一步来说，要想给出 $M_2(D)$ 与D的一般的映射，倒并不困难．比如，规则

$$\psi:(a_{ij}) \longmapsto a_{11},$$

即规则 ψ把二阶方阵 (a_{ij})射成它的左上角的那个元素 a_{11}。显然这个规则 ψ是从 $M_2(D)$ 到D的一个映射，而且是满映射，但不是单映射。

我们应当提出，这个映射 ψ说明了 $M_2(D)$ 与D之间的什么问题、体现着什么联系？当然，我们不能说映射 ψ不说明任何问题，也没体现任何联系。不过应该承认，映射 ψ仅仅说明了这样的问题和联系：每一二阶方阵 (a_{ij}) 的元素都是D中的一个确定的数、D中的任何一个数都能够做为某一二阶方阵 (a_{ij}) 左上角的元素、但是不同的两个二阶方阵左上角的元素可能相同．诚然，ψ 所说明的这种联系是平凡的不足道的。但是，它毕竟还是说明了一定的、实在的问题和联系．可是让我们再指出以下的，也不宜忽视的情况．我们给出一个规则如下：

$$\varphi:\ (a_{ij}) \longmapsto 1.$$

不能不承认，规则 φ 是 $M_2(D)$与D的一个映射，它把每一个二阶方阵都射成１．试问，映射 φ说明什么问题，又体现什么联系呢？很难答，也实在说不清楚．因为没法解释，为什么每一个二阶方阵都与１有必然的联系。对此我们应当这样看，从数学的意义上说，

我们不否认 φ是映射，但是它说明什么问题，体现什么联系，与我们对于所论问题认识的深度与广度有关系，同时也依赖于讨论问题的背景。

§3 代数运算

代数运算的概念是清楚的，没什么费解之处，只是有一点明确一下为好。即，所说集A的运算就是$A\times A$到A的映射。因此，如果 $a,b\in A$ 但 $a\neq b$，那么 (a,b)与 (b,a) 都是$A\times A$的元素，但它们是不同的两个元素，即 $(a,b)\neq(b,a)$。从而在 $A\times A$ 到A的映射φ之下的象不必相同。换句话说就是，$\varphi((a,b))=\varphi((b,a))$即 $a\varphi b=b\varphi a$不一定成立。只有 φ满足交换律时才有对任意的$a,b\in A$，$a\varphi b=b\varphi a$ 成立。

关于结合律、交换律、分配律有相应的三个命题。它们的证明方法都是很明白的。只是在推导的过程中应当弄清楚每一个步骤的依据。

下面我们把关于结合律的命题 1 的证明中的推导过程重写于下，把每一步骤的依据用括号标注于后。

$$
\begin{aligned}
\pi(a_1\circ\cdots\circ a_n) &= \pi(a_1\circ\cdots\circ a_i)\circ\pi(a_{i+1}\circ\cdots\circ a_n) && \text{(两得一个)}\\
&= \pi(a_1\circ\cdots\circ a_i)\circ(\pi(a_{i+1}\circ\cdots\circ a_{n-1})\circ a_n) && \text{(归纳假设)}\\
&= (\pi(a_1\circ\cdots\circ a_i)\circ\pi(a_{i+1}\circ\cdots\circ a_{n-1}))\circ a_n && \text{(结合律)}\\
&= \pi(a_1\circ\cdots\circ a_{n-1})\circ a_n && \text{(归纳假设)}\\
&= ((\cdots((a_1\circ a_2)\circ a_3)\cdots)\circ a_{n-1})\circ a_n && \text{(归纳假设)}
\end{aligned}
$$

再强调一下：代数运算是一个非常一般的概念。它来源于数的四则运算，但切不可与数的四则运算等同起来。事实上，凡是“两得一个”的方法都是代数运算。这在正文中已有所论述，此处就不再多说什么。

下面我们举一个重要的例子。

设A，B，C 为任意三个非空集合。用 $H(A,B)$表示从A到B

的一切映射的集合。同样地，用 $H(B,C)$，$H(A,C)$ 分别表示从 B 到 C 的一切映射的集合与从 A 到 C 的一切映射的集合。

任取 $\psi\in H(B,C)$，$\varphi\in H(A,B)$。于是对任一元素 $a\in A$，都有

$$\varphi: a \longmapsto \varphi(a)\in B,\quad \psi: \varphi(a) \longmapsto \psi(\varphi(a))\in C.$$

如果规定

$$(\psi,\varphi): a \longmapsto \psi(\varphi(a)),$$

那么这个规则 (ψ,φ) 就是从 A 到 C 的一个映射，即 $(\psi,\varphi)\in H(A,C)$。我们把 A 到 C 的这个映射叫做 ψ 与 φ 的合成，记作，$(\psi,\varphi)=\psi\varphi$，即

$$(\psi\varphi)(a)=\psi(\varphi(a)),\ \forall a\in A.$$

用上述的"合成方法"，从任意一个 $\psi\in H(B,C)$ 和任意一个 $\varphi\in H(A,B)$ 都能得到一个映射 $\psi\varphi\in H(A,C)$。这样，规则

$$0:\ (\psi,\varphi) \longmapsto \psi\varphi$$

就是 $H(B,C)\times H(A,B)$ 到 $H(A,C)$ 的一个映射，亦即 $H(B,C)$ 与 $H(A,B)$ 到 $H(A,C)$ 的一个代数运算。象这样把两个映射合成一个映射的方法，通常就叫做映射的乘法。合成的映射 $\psi\varphi$ 叫做 ψ 与 φ 的积。

特别重要的是 $A=B=C$ 的特殊情形。这时所有的映射 ψ，φ，$\psi\varphi$ 全都是集合 A 的变换。比如取 $A=D$。于是 $H(D,D)$ 就是以实数集 D 为定义域的一切函数的集合。任取 $f,g\in H(D,D)$，则有

$$(fg)(x)=f(g(x)),\ \forall x\in D.$$

因此，这时所说的 f 与 g 的合成函数 fg，正是通常的 f 与 g 的复合函数。所以，一般的变换合成或变换乘法就是通常复合函数概念的推广。

§4 集合的子集分解

学习本节应当注意以下四个问题：

1．弄通集合的分类的定义；

2．做成分类的一般方法；

3．分类、等价关系的例子；

4．用分类的观念说明一些具体问题．

第一点是简单的．因为分类的定义本身并不复杂，而且有很大的直观性．

第二点所说分类的一般方法就是利用等价关系进行分类的方法．等价关系是数学中极其广泛的概念，有很大的概括性，是相当重要的一种数学观念．再有，等价关系的定义本身也是比较复杂的．我们认为，说明等价关系与分类之间的联系的本节定理是掌握等价关系的实质的重要线索．同时，对数学中“等价”的含义必须从数学的角度领会它的具体的针对性，即每一具体条件下的“等价”都有它具体的特殊的含义，以免笼统的空洞的“理解”．

例如，对于整数的整除问题，把互相整除的整数叫做是等价的，因为二者在整除的问题中有着完全相同的作用．

对于多项式的整除问题，也是把互相整除的多项式叫做是等价的，因为二者在整除的问题中有着完全相同的作用．

在线性方程组的问题中，把同解的线性方程组叫做是等价的，因为二者在求解的问题中有着完全相同的作用．

对于用分离系数法解线性方程组的问题，把行相抵的矩阵叫做是等价的，因为二者在线性方程组的求解问题中有着完全相同的作用．

对n维向量的线性关系问题，把能互相线性表示的向量组叫做是等价的，因为二者在向量的线性表示的问题中有着完全相同的作用．

对矩阵的秩数问题，把相抵矩阵叫做是等价的，因为二者在求秩的问题中有着完全相同的作用．

掌握一批恰当的例子，这对具体理解分类、等价关系的概念和方法，无疑是十分重要的．正文中的例子已经比较充分，应当首先

熟悉这些例子。

用分类或等价关系的方法说明某些问题，关键在于对所讨论的对象有某种规律性的认识，不然，分类就无所依据和无的放矢，或者说无从发现应把什么关系视为等价。

〔例题选解〕

例 1　证明：$(A-B)\cup(B-A)=(A\cup B)-(A\cap B)$。

证明　设 $x\in(A-B)\cup(B-A)$，则有$x\in A-B$或$x\in B-A$。若 $x\in A-B$，即 $x\in A$ 但 $x\overline{\in}B$。从而 $x\in A\cup B$，$x\overline{\in}A\cap B$，所以 $x\in(A\cup B)-(A\cap B)$。类似地，若 $x\in B-A$，即 $x\in B$ 但 $x\overline{\in}A$，从而 $x\in A\cup B$，$x\overline{\in}A\cap B$，所以 $x\in(A\cup B)-(A\cap B)$。这就得出 $(A-B)\cup(B-A)\subseteq(A\cup B)-(A\cap B)$。

另一方面，设 $x\in(A\cup B)-(A\cap B)$，则有 $x\in A\cup B$ 但 $x\overline{\in}A\cap B$。从而 $x\in A$ 或 $x\in B$ 但 $x\overline{\in}A\cap B$。亦即

$$x\in A \text{ 但 } x\overline{\in}A\cap B \text{ 或 } x\in B \text{但 } x\overline{\in}A\cap B,$$

然而 $x\in A$，$x\overline{\in}A\cap B$ 就是 $x\in A$ 但 $x\overline{\in}B$ 即 $x\in A-B$，同理，$x\in B$ $x\overline{\in}A\cap B$ 就是 $x\in B$ 但 $x\overline{\in}A$ 即 $x\in B-A$。这样就得$x\in A-B$ 或 $x\in B-A$，亦即 $x\in(A-B)\cup(B-A)$。从而 $(A\cup B)-(A\cap B)\subseteq(A-B)\cup(B-A)$。总之则得

$$(A-B)\cup(B-A)=(A\cup B)-(A\cap B).$$

例 2　令 S_n 表示 1，2，…，n这n个数码的一切全排列组成的集合。考虑 S_n 到整数集 Z 的映射。为了建立 S_n 与 Z 之间的映射，就是要指出 n元排列与整数之间的某种联系。这是可能办到的，因为我们已经了解每一 n元排列都与一个确定的非负整数有关系，即它的反序数。这样，把 n元排列与其反序数之间的联系表达出来就给出一个从 S_n 到 Z 的映射。具体的规则为

$$\varphi:\ i_1 i_2 \cdots i_n \longmapsto \tau(i_1\ i_2\ \cdots\ i_n).$$

显然，φ 既不是单映射也不是满映射.

由映射 φ给 S_n 带来一个分类，即

$i_1\ i_2\ \cdots\ i_n$ 与 $j_1\ j_2\ \cdots\ j_n$ 分在一类当且仅当

$\tau(i_1\ i_2\ \cdots\ i_n)=\tau(j_1\ j_2\ \cdots\ j_n)$.

换句话说就是，如果 $i_1\ i_2\ \cdots\ i_n$ 与 $j_1\ j_2\ \cdots\ j_n$ 的反序数相等，那么 $i_1\ i_2\ \cdots\ i_n$ 与 $j_1\ j_2\ \cdots\ j_n$ 就分在一类里，否则就不分在一类里.

我们以 $n=3$ 为例，写出这样的具体分类.

$S_3=\{123,\ 132,\ 213,\ 231,\ 312,\ 321\}$.

$$\begin{aligned}\varphi:\ &123 \longmapsto 0,\\ &132 \longmapsto 1,\\ &213 \longmapsto 1,\\ &231 \longmapsto 2,\\ &312 \longmapsto 2,\\ &321 \longmapsto 3.\end{aligned}$$

于是以下的四个子集：

$A_0=\{123\}$, $A_1=\{132, 213\}$, $A_2=\{231, 312\}$, $A_3=\{321\}$

就构成 S_3 的一个分类 $\Sigma=\{A_0,\ A_1,\ A_2,\ A_3\}$.

我们知道，在行列式的理论中，关于排列，需要的仅仅是它们的奇偶性，反序数的具体数值是不关紧要的. 这个事实可以通过以下的映射得到表达. 即规则

$$\begin{aligned}\psi:\ &i_1\ i_2\ \cdots\ i_n \longmapsto 0,\ \text{如果}\ 2\mid\tau(i_1\ i_2\ \cdots\ i_n),\\ &i_1\ i_2\ \cdots\ i_n \longmapsto 1,\ \text{如果}\ 2\nmid\tau(i_1\ i_2\ \cdots\ i_n).\end{aligned}$$

是 S_n 到 Z 的一个映射，而这个映射给 S_n 带来的分类就把 n 元排列分成两类：偶排列组成一类，奇排列组成一类. 即

$$\overline{\Sigma}=\{B_0,\ B_1\},$$

其中 $B_0=\{i_1\ i_2\ \cdots\ i_n \mid 2\mid\tau(i_1\ i_2\ \cdots\ i_n)\}$, $B_1=\{i_1\ i_2\ \cdots\ i_n \mid 2\nmid\tau(i_1\ i_2\ \cdots\ i_n)\}$.

例3　考虑 $F[x]$，正整数集N和非负整数集N_0。我们给出一个规则

$$\varphi: f(x) \longmapsto \deg f(x),$$

即这个规则 φ 把 $f(x)$ 射成它的次数 $\deg f(x)$。我们问，把这个规则 φ 用于 $F[x]$ 与正整数集N之间时，它是不是一个映射？用于 $F[x]$与 N_0 之间又如何？

N 是正整数集，而 $F[x]$ 中有一些多项式的次数等于0。这样的多项式规则 φ都没能在N中给规定出象来。因此规则φ不是$F[x]$与N之间的映射。非负整数集N_0中含有0，这样，把规则 φ 用于 $F[x]$ 与 N_0 之间时，对于 $F[x]$ 中每一零次多项式，规则 φ 把它们都射成 N_0 中的0，即在 φ 之下，零次多项式在 N_0 中都有象。尽管如此，规则 φ 还不是从 $F[x]$ 到 N_0 的映射。这是因为 $F[x]$ 中有一个特殊的没有次数的多项式，即零多项式0。对于这个没有次数的多项式来说，规则 φ当然不可能给它规定出象来。即规则 φ不是对 $F[x]$ 中每一个元素都起作用。所以规则 φ 也不是 $F[x]$ 与 N_0 之间的映射。

次数是多项式的一个很重要的特征。仅仅因为这一个特殊的没有次数的零多项式，在考虑多项式与其次数的联系时，显得不大方便，于是，如能给零多项式安排一个恰当的“次数”是有意义的。出自这种考虑，根据经验把零多项式的次数记作$-\infty$，并且规定对任意非负整数n，都有

$$-\infty < n,$$

$$-\infty + n = n + (-\infty), \quad -\infty + (-\infty) = -\infty.$$

在这样的安排下，把符号$-\infty$放到 N_0 里去，即令

$$N_1 = \{-\infty, \ 0, 1, 2, 3, \ \cdots\}.$$

于是，把规则 φ用于 $F[x]$ 与 N_1 之间时，它就是从$F[x]$到N_1的一个映射，而且是满映射。这时当然有 $\varphi(0) = -\infty$。

我们说，给零多项式如上那样安排的次数是恰当的、合理的。因为关于次数的两条重要性质，在这样的安排下仍然成立。即

$$\deg(f(x)+g(x))\leqslant\max(\deg f(x),\ \deg g(x)),$$
$$\deg(f(x)g(x))=\deg f(x)+\deg g(x).$$

这个映射 φ 给 $F[x]$ 带来的分类是：

$f(x)$ 与 $g(x)$ 分在一类当且仅当 $\deg f(x)=\deg g(x)$。换句话说就是，如果 $f(x)$ 与 $g(x)$ 的次数相同，那么 $f(x)$ 与 $g(x)$ 就分在一类里，否则就不分在一类里。这个分类就是 §4 例 4 中的分类：

$$\Sigma=\{\{0\},\ d_0,\ d_1,\ d_2,\ \cdots\}.$$

如果按照符号： $d_i=\{f(x)\mid \deg f(x)=i\}$ 的写法，零多项式组成的类{ 0 }就可以写成 $d_{-\infty}=\{0\}$。这样就是 $\Sigma=\{d_{-\infty}, d_0, d_1, d_2, \cdots\}$。

这个分类所决定的等价关系～是：

$f(x)\sim g(x)$ 当且仅当 $f(x)$ 与 $g(x)$ 分在一类里。换个说法就是：

$f(x)\sim g(x)$ 当且仅当 $\deg f(x)=\deg g(x)$。

$F[x]$ 有一个熟知的等价关系，即互相整除关系。显然，互相整除的两个多项式它们的次数必相等，但反之则未必对。这说明次数相等的关系与互相整除的关系不完全一致。

例 4　S_n 为 n 元排列的集合。我们回忆一下排列的对换的概念。即，对换就是互换一个排列中某两个数码的位置的一种方法。它把每一个 n 元排列变成另一个确定的 n 元排列。为了行文方便，我们约定，如果排列 $i_1\ i_2\ \cdots\ i_n$ 经过一次对换 (s,t) 之后得到的排列为 $j_1\ j_2\ \cdots\ j_n$，那么就记作 $i_1\ i_2\ \cdots\ i_n^{(s,t)}=j_1\ j_2\ \cdots\ j_n$。比如

$$3142^{(3,4)}=4132,\quad 12345^{(5,2)}=15342.$$

这样，如果把对换 (s,t) 做为一个规则，即

$$(s,t):i_1\ i_2\ \cdots\ i_n\longmapsto i_1\ i_2\ \cdots\ i_n^{(s,t)},$$

那么 (s,t) 就是 S_n 的一个变换。容易看出，(s,t) 是 S_n 的单满变换，即一一变换。从而 (s,t) 有逆变换，并且 (s,t) 的逆变换就是 (s,t) 本身。

例 5　考虑实数集 D，给出以下两条规则：

$$\varphi_1: (a,b) \longmapsto \max(a,b),$$

$$\varphi_2: (a,b) \longmapsto \min(a,b).$$

此处的规则 φ_1 就是把 (a,b) 射成 a 与 b 中较大的数；而规则 φ_2 把 (a,b) 射成 a与 b 中较小的数．比如

$$\varphi_1((2,7)) = 2\varphi_1 7 = \max(2,7) = 7,$$

$$\varphi_1((-5,0)) = -5\varphi_1 0 = \max(-5,0) = 0,$$

$$\varphi_1((-2,-2)) = -2\varphi_1(-2) = \max(-2,-2) = -2,$$

$$\varphi_2((2,7)) = 2\varphi_2 7 = \min(2,7) = 2,$$

$$\varphi_2((-5,0)) = -5\varphi_2 0 = \min(-5,0) = -5,$$

$$\varphi_2((-2,-2)) = -2\varphi_2(-2) = \min(-2,-2) = -2.$$

显然，φ_1 与 φ_2 都是实数集 D的代数运算．我们容易指出，φ_1 与 φ_2都满足结合律同时也都满足交换律．事实上，$\forall a, b, c \in D$，则有

$$\begin{aligned}(a\varphi_1 b)\varphi_1 c &= \max(a,b)\varphi_1 c = \max(\max(a,b),c)\\ &= \max(a, b, c),\\ a\varphi_1(b\varphi_1 c) &= a\varphi_1 \max(b,c) = \max(a, \max(b,c))\\ &= \max(a, b, c).\end{aligned}$$

所以

$$(a\varphi_1 b)\varphi_1 c = a\varphi_1(b\varphi_1 c).$$

类似地，可得 $(a\varphi_2 b)\varphi_2 c = a\varphi_2(b\varphi_2 c)$.

交换律是自明的．

下面证明这两个运算之间也适合分配律，即有

$$a\varphi_1(b\varphi_2 c) = (a\varphi_1 b)\varphi_2(a\varphi_1 c) \tag{1}$$

$$a\varphi_2(b\varphi_1 c) = (a\varphi_2 b)\varphi_1(a\varphi_2 c) \tag{2}$$

我们分几种情形验证（1）式成立．

1）当 $a \leqslant b \leqslant c$ 时，有

$$a\varphi_1(b\varphi_2 c) = a\varphi_1 b = b, \quad (a\varphi_1 b)\varphi_2(a\varphi_1 c) = b\varphi_2 c = b,$$

所以 $a\varphi_1(b\varphi_2 c) = (a\varphi_1 b)\varphi_2(a\varphi_1 c)$.

2）当 $b \leqslant a \leqslant c$ 时，有

$$a\varphi_1(b\varphi_2c)=a\varphi_1b=a,\quad (a\varphi_1b)\varphi_2(a\varphi_1c)=a\varphi_2c=a,$$

所以 $a\varphi_1(b\varphi_2c)=(a\varphi_1b)\varphi_2(a\varphi_1c)$。

3）当 $b\leqslant c\leqslant a$ 时，有

$$a\varphi_1(b\varphi_2c)=a\varphi_1b=a,\quad (a\varphi_1b)\varphi_2(a\varphi_1c)=a\varphi_2a=a,$$

所以 $a\varphi_1(b\varphi_2c)=(a\varphi_1b)\varphi_2(a\varphi_1c)$。

对于 $a\leqslant c\leqslant b$，$c\leqslant a\leqslant b$，$c\leqslant b\leqslant a$ 这三种情形，利用交换律可以相应地归为以上三种情形。这就证明了，对于任意的 $a, b, c\in D$，（1）式成立。

完全类似地，可证（2）式成立。

例 6　导出映射

设 φ是从A到 B 的映射。S 是A的一个非空子集。于是如果仅仅考虑 φ 在子集 S 上的作用，那么规则 φ 自然也是从 S 到 B 的映射：

$$\varphi:\ S\longrightarrow B.$$

我们称 S 的这个映射为 φ在 S 上的**导出映射**，记作 $\varphi|_S$。特别地，当 $A=B$ 时，如果子集 S 具有性质：$\forall x\in S$，$\varphi(x)\in S$，简记作 $\varphi(S)\subseteq S$，那么φ 作用在 S 上时，就是 S 的一个变换，称其为 φ 在 S 上的**导出变换**。

比如，整数集 Z 到非负整数集 N_0 的映射

$$\varphi:\ n\longmapsto |n|$$

在 N_0 上的导出映射为

$$\varphi|_{N_0}:\ m\longmapsto |m|=m,\ \forall m\in N_0.$$

而且这个导出映射是 N_0 的单满映射。

再如，$M_2(D)$ 到 $B=\{0,1,2\}$ 的映射

$$\varphi:\ (a_{ij})\longmapsto \mathrm{rank}\,(a_{ij})$$

在子集 $S_1=\{(a_{ij})\mid \det(a_{ij})\neq 0\}$ 上的导出映射为

$$\varphi|_{S_1}:\ (a_{ij})\longmapsto 2,\ \forall (a_{ij})\in S_1;$$

在子集 $S_2=\left\{\begin{pmatrix}0&0\\0&0\end{pmatrix},\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}1&0\\0&1\end{pmatrix}\right\}$ 上的导出映射为

$$\varphi|S_2:\begin{pmatrix}0&0\\0&0\end{pmatrix}\longmapsto 0,\quad \begin{pmatrix}1&0\\0&0\end{pmatrix}\longmapsto 1,\quad \begin{pmatrix}1&0\\0&1\end{pmatrix}\longmapsto 2.$$

显然，$\varphi|S_2$ 是 S_2 与 B 之间的单满映射，即一一映射。

又如，$D[x]$ 的变换

$$\varphi: f(x)\longmapsto f'(x)$$

在子集 $D_n[x]=\{f(x)|\deg f(x)<n\}$ 上可以导出一个变换。因为对任一 $f(x)\in D_n[x]$，$\varphi(f(x))=f'(x)\in D_n[x]$。

又如，$M_{m,n}(F)$ 的任一初等变换都不变矩阵的秩。从而每一个初等变换：$D_i(\lambda)$，$P_{ij}(k)$，$C(i,j)$，都在子集 $d_r=\{(a_{ij})|\operatorname{rank}(a_{ij})=r\}$ 上导出一个变换。

第八章　矩阵的运算

〔内容提要〕

这一章主要讲矩阵的三种运算，重点围绕矩阵乘法运算而展开的。在第一节讲了矩阵的加法，倍数乘法和乘法运算及其性质之后；在第二节从矩阵乘法出发引进了可逆矩阵的概念，给出了可逆矩阵的判别法与逆矩阵的求法；在第三节引进了特殊的可逆矩阵——初等矩阵，通过初等矩阵与初等变换之关系的讨论，给出了可逆矩阵与初等矩阵之关系以及证明了乘积阵的行列式等于各因子的行列式之积；乘积阵的秩不超过每个因子阵的秩；两个矩阵相抵的充要条件等重要定理。

本章的主要概念是：

矩阵的加法、倍数乘法及乘法运算的定义，可逆矩阵，初等矩阵等概念。

本章的主要结果是：

1　矩阵的加法、倍数乘法及乘法运算的 1）——11）条性质。

2　n 阶方阵A可逆的充分必要条件：

1）A可逆的充分必要条件是：$|A|\neq 0$；

2）A可逆的充分必要条件是：$A=P_1P_2\cdots P_s$，其中 P_i为初等矩阵（$i=1,2,\cdots,s$）；

3）A可逆的充分必要条件是：

$$A=P_1P_2\cdots P_s\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix},$$

其中 P_i 都是消法矩阵，$d_j \neq 0$，$i=1,2,\cdots,s$；$j=1,2,\cdots,n$。

A可逆的充分必要条件是：

$$A = P_1 P_2 \cdots P_t \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d \end{pmatrix},$$

其中 P_i 都是消法矩阵，$d=\det A$，$(i=1,2,\cdots,t)$。

3　A与B相抵的充分必要条件是：存在可逆矩阵P，Q，使得

$$B = PAQ.$$

进而有：$\operatorname{rank}A=r$ 的充分必要条件是：存在可逆矩阵P,Q，使得

$$A = P\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}Q.$$

（其中1共 r 个）

4　$\det AB = \det A \cdot \det B$

5　$\operatorname{rank}AB = \operatorname{rank}B$，当$A$为可逆的；

$\operatorname{rank}AB \leqslant \min(\operatorname{rank}A, \operatorname{rank}B)$。

本章的主要方法是：

1．判别 n阶方阵A是否可逆的方法主要有以下三种：

1）直接按定义：即看是否有 n阶方阵B存在，使得

$AB=BA=I_n$；

2）看A的行列式 $\det A$ 是否为0；

3）看A在相抵之下的标准形是否为单位矩阵。

2　求可逆矩阵A的逆矩阵的方法：

1）利用公式

$$A^{-1}=\frac{1}{|A|}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix},$$

其中 A_{ij} 为 $A=(a_{ij})$ 中元素 a_{ij} 的代数余子式；

2）初等变换求逆阵的方法，即

$(A\ I_n)\rightleftarrows(I_n\ A^{-1})$。

通过这一章的学习，要求读者

1）正确理解和掌握本章的主要概念，熟练而准确地进行矩阵（包括分块阵）的运算。

2）会判别一个方阵是否可逆，并会求逆矩阵。

3）掌握初等变换与初等矩阵，可逆矩阵与初等矩阵之关系。

4）能利用本章主要结果论证有关问题。

〔内容分析〕

§1 矩阵的运算及其性质

本节主要讲矩阵的三种运算及其性质：

加法运算及其性质1）—4），同时利用加法定义了减法；倍数乘法运算及其性质5）—8）；矩阵乘法及其性质9）—11）。在此基础上介绍了分块矩阵的运算及转置阵的性质。

通过这一节的学习，要求读者熟练地掌握矩阵的三种运算及其性质。

在学习这节时需要注意的是：

1）对给定的两个矩阵A与B，存在可不可加，可不可乘的问题。只有当A的行数与B的行数相等，A的列数与B的列数也相等才可以相加减；只有A的列数等于B的行数才可相乘。

2）分块矩阵的加法与乘法也存在可不可运算的问题．只有A的行数与B的行数相等，A的列数与B的列数相等，同时A的分块方法与B的分块方法完全相同时才可以相加减；只有A的列数等于B的行数，A的列的分法与B的行的分法相同时才可以相乘．

3）在矩阵运算中，有一些数的运算性质在矩阵中是不成立的．比如

$$A=\begin{pmatrix}1&0\\0&0\end{pmatrix}\neq 0,\quad B=\begin{pmatrix}0&0\\0&1\end{pmatrix}\neq 0,\quad AB=\begin{pmatrix}0&0\\0&0\end{pmatrix}=0;$$

$$A=\begin{pmatrix}0&1\\0&0\end{pmatrix}\neq 0,\quad A^2=AA=\begin{pmatrix}0&1\\0&0\end{pmatrix}\begin{pmatrix}0&1\\0&0\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix}=0.$$

上二式说明：A,B 都是非零矩阵，但是乘积却为零阵；A为非零矩阵，但A的若干次幂等于零阵．还有正文中已指出的乘法不满足交换律等．所有这些告诉我们，在数的运算中成立的性质，在矩阵中在没有证明成立之前是不能应用的．

§2 可逆矩阵

本节主要讲什么是可逆矩阵，如何判别一个矩阵是否是可逆阵，以及如何求一个矩阵的逆矩阵等三个问题．另外还讨论了可逆矩阵的简单性质和矩阵方程的求解问题．

通过这一节的学习，要求读者正确地掌握可逆矩阵的定义，会判别一个矩阵是否为可逆矩阵，能够准确地求一个矩阵的逆阵．

§3 初等矩阵

本节主要讲三个方面的问题：

1）讲初等矩阵与初等变换之关系，即命题；

2）讲可逆矩阵与初等矩阵之关系，即定理1、定理2及其推

论；

3）给出了三个重要结论：A与B相抵的充分必要条件，即定理3；乘积阵的行列式等于各因子阵行列式之积，即定理4；乘积阵的秩不超过每个因子阵的秩，即定理5，定理6。

上述三个问题之关系是：由1）推出2），由1），2）推出3）。在2）中给出了一个求逆阵的重要方法，即初等变换法。

通过这一节的学习，要求读者

1）掌握初等矩阵与初等变换之关系，可逆矩阵与初等矩阵之关系。

2）会利用初等变换法求逆阵。

3）会利用本节的主要结论论证有关问题。

〔例题选解〕

例1　如果 $AB=BA$，则说A与B可交换。求所有与

$$A=\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$$

可交换的矩阵。

解　设

$$X=\begin{pmatrix}x_1&y_1&z_1\\x_2&y_2&z_2\\x_3&y_3&z_3\end{pmatrix}.$$

那么

$$AX=\begin{pmatrix}x_1+x_2&y_1+y_2&z_1+z_2\\x_2+x_3&y_2+y_3&z_2+z_3\\x_3&y_3&z_3\end{pmatrix}$$

$$XA=\begin{pmatrix} x_1 & x_1+y_1 & y_1+z_1 \\ x_2 & x_2+y_2 & y_2+z_2 \\ x_3 & x_3+y_3 & y_3+z_3 \end{pmatrix}.$$

因此，X与A可交换的充分必要条件是

$$\begin{cases} x_1+x_2=x_1 \\ y_1+y_2=x_1+y_1 \\ z_1+z_2=y_1+z_1 \\ x_2+x_3=x_2 \\ y_2+y_3=x_2+y_2 \\ z_2+z_3=y_2+z_2 \\ y_3=x_3+y_3 \\ z_3=y_3+z_3. \end{cases}$$

而这组条件又等价于

$$\begin{cases} x_2=x_3=y_3=0 \\ x_1=y_2=z_3 \\ y_1=z_2. \end{cases}$$

所以与A可交换的全部矩阵都可表成

$$X=\begin{pmatrix} x_1 & y_1 & z_1 \\ 0 & x_1 & y_1 \\ 0 & 0 & x_1 \end{pmatrix},$$

其中 x_1, y_1, z_1 可以取数域 F中任意的数。

例2 设 A,B 为 n 阶方阵．试问：在什么条件下，有

$$(A+B)^2=A^2+2AB+B^2 \tag{1}$$

成立？

解 首先，我们看当（1）成立时 A,B 应适合什么条件．因为

$$(A+B)^2=(A+B)(A+B)=A^2+AB+BA+B^2.$$

若（1）成立，则有

$$A^2+2AB+B^2=A^2+AB+BA+B^2.$$

故必有

$$BA=AB.$$

其次，当 $AB=BA$ 时显然（1）式成立。于是（1）成立的充分必要条件是：

$$AB=BA.$$

例 3 设 $A=\frac{1}{2}(B+I_n)$。证明 $A^2=A$ 的充分必要条件是：$B^2=I_n$。

证明 必要性。因为

$$A^2=\left[\frac{1}{2}(B+I_n)\right]^2=\frac{1}{4}(B+I_n)^2=\frac{1}{4}(B^2+2B+I_n).$$

由于 $A^2=A$，所以有

$$\frac{1}{4}(B^2+2B+I_n)=\frac{1}{2}(B+I_n).$$

即

$$\frac{1}{4}B^2+\frac{1}{4}I_n=\frac{1}{2}I_n.$$

因此 $$B^2=I_n.$$

充分性。因为 $B^2=I_n$，所以有

$$A^2=\left(\frac{1}{2}(B+I_n)\right)^2=\frac{1}{4}(B^2+2B+I_n)=\frac{1}{4}(2B+2I_n)$$

$$=\frac{1}{2}(B+I_n)=A.$$

即

$$A^2=A.$$

例 4 设 A,B 都是 n 阶可逆矩阵，则

$$D=\begin{pmatrix}A & 0\\ C & B\end{pmatrix}$$

必为可逆矩阵，并求 D 的逆阵。

证明　因为

$$\det D = \det A \cdot \det B \neq 0,$$

故 D 为可逆矩阵。

设

$$D^{-1} = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}.$$

由于

$$DD^{-1} = \begin{pmatrix} A & 0 \\ C & B \end{pmatrix} \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}$$

$$= \begin{pmatrix} AX_{11} & AX_{12} \\ CX_{11} + BX_{21} & CX_{12} + BX_{22} \end{pmatrix}$$

$$= \begin{pmatrix} I_n & 0 \\ 0 & I_n \end{pmatrix}.$$

根据矩阵相等的定义，有

$$\begin{cases} AX_{11} = I_n \\ AX_{12} = 0 \\ CX_{11} + BX_{21} = 0 \\ CX_{12} + BX_{22} = I_n. \end{cases}$$

由前二式知，$X_{11} = A^{-1}$，$X_{12} = 0$。代入后二式，得

$$X_{22} = B^{-1},\ X_{21} = -B^{-1}CA^{-1}.$$

故

$$D^{-1} = \begin{pmatrix} A^{-1} & 0 \\ -B^{-1}CA^{-1} & B^{-1} \end{pmatrix}.$$

例 5　设 A 为 n 阶方阵，且 $A \neq 0$。证明：存在非零的 n 阶方阵 B，使得

$$AB = 0 \tag{2}$$

的充分必要条件是：$|A| = 0$。

证明　必要性。由 $AB = 0$，去证 $|A| = 0$。反证，假定 $|A| \neq 0$，

则A为可逆矩阵，令A的逆阵为A^{-1}。用A^{-1}左乘（2）式两端，有

$$A^{-1}(AB)=A^{-1}0,\quad 即\quad B=0,$$

这与B非零矛盾，故必有$|A|=0$。

充分性。若$|A|=0$，去证存在非零阵B，使$AB=0$。考虑线性方程组

$$A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}.$$

由于系数阵A的行列式$|A|=0$，必有非零解，取一个非零解

$$\begin{pmatrix}x_{10}\\ \vdots\\ x_{n0}\end{pmatrix}$$

做为B阵的第一列，其余列都取0，即

$$B=\begin{pmatrix}x_{10} & 0 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ x_{n0} & 0 & \cdots & 0\end{pmatrix}\neq 0$$

而$AB=0$。

例6　设A,B为二n阶方阵。证明：若$AB=0$，则$\operatorname{rank}A+\operatorname{rank}B\leqslant n$。

证明　设$\operatorname{rank}A=r$，$B=(B_1B_2\ \cdots\ B_n)$。则

$$AB=A(B_1B_2\ \cdots\ B_n)=0.$$

于是

$$AB_1=0,\quad AB_2=0,\quad \cdots,\quad AB_n=0.$$

这说明：B阵的各个列B_1，B_2，…，B_n都是以A为系数阵的齐次线性方程组

$$AX=0 \tag{3}$$

的解。由于$\operatorname{rank}A=r$，故（3）的解空间的维数为$n-r$。从而可知列向量组B_1，B_2，…，B_n的极大无关组所含向量的个数不超过

$n-r$。因此，$\operatorname{rank}B\leqslant n-r$。故

$$\operatorname{rank}A+\operatorname{rank}B\leqslant r+n-r=n.$$

例 7　证明

$$|A^*|=|A|^{n-1},$$

其中 A 为 n 阶方阵（$n\geqslant 2$）。

证明　因为

$$AA^*=\begin{pmatrix}|A| & & \\ & \ddots & \\ & & |A|\end{pmatrix}.$$

所以

$$|AA^*|=|A|^n,\ \text{即}\ |A|\,|A^*|=|A|^n.$$

当 $|A|\neq 0$ 时，有

$$|A^*|=|A|^{n-1};$$

当 $|A|=0$ 时，一种情形：$\operatorname{rank}A=0$，即 $A=0$，这时 $A^*=0$。故有

$$|A^*|=|A|^{n-1};$$

另一种情形：$0<\operatorname{rank}A<n$。由于

$$AA^*=\begin{pmatrix}|A| & & \\ & \ddots & \\ & & |A|\end{pmatrix}=0,$$

据上例可知，有

$$\operatorname{rank}A+\operatorname{rank}A^*\leqslant n.$$

又因 $\operatorname{rank}A>0$，所以必有

$$\operatorname{rank}A^*<n,\ \text{从而}\ |A^*|=0.$$

故也有

$$|A^*|=|A|^{n-1}.$$

例 8　设 A 为 $l\times m$ 矩阵，$\operatorname{rank}A=r_A$，B 为 $m\times n$ 矩阵，$\operatorname{rank}B=r_B$。证明

$$\operatorname{rank}AB\geqslant r_A+r_B-m.$$

证明　由 $\operatorname{rank} B=r_B$，据本章定理3的推论知，存在$m$阶可逆矩阵$P$，$n$阶可逆矩阵$Q$，使得

$$B=P\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}Q.$$

于是

$$AB=AP\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}Q,$$

$$\operatorname{rank} AB=\operatorname{rank}\left[AP\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}Q\right]$$

$$=\operatorname{rank}\left[AP\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}\right].$$

而一方面，由于

$$\operatorname{rank} AP=\operatorname{rank} A=r_A.$$

令

$$AP=(A_1\ \ A_2),$$

其中A_1为$l\times r_B$矩阵，A_2为$l\times(m-r_B)$矩阵。根据第六章知，有

$$\operatorname{rank} A_1\geqslant r_A-(m-r_B)=r_A+r_B-m.$$

另一方面

$$AP\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}=(A_1\ \ A_2)\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix}.$$

因此

$$\operatorname{rank}\left[AP\begin{pmatrix} I_{r_B} & 0 \\ 0 & 0 \end{pmatrix}\right]=\operatorname{rank}\begin{pmatrix} A_1 & 0 \\ 0 & 0 \end{pmatrix}=\operatorname{rank} A_1$$

$$\geqslant r_A+r_B-m.$$

故

$$\operatorname{rank} AB\geqslant r_A+r_B-m.$$

第九章　对称阵在相合之下的标准形

〔内容提要〕

这一章主要是解决对称阵在相合之下的标准形问题。第一节指出了讨论对称阵在相合之下的标准形的背景和相合分类。第二节证明了对称阵都相合于对角形阵，且对角线上是 1 或 -1 或 0，即标准形。同时给出了两种求标准形的方法，即初等变换法和配方法。接着在第三节证明了对称阵标准形的唯一性。第四节进一步讨论了特殊的对称矩阵，即正定矩阵及其性质。

本章的主要概念是：

二次型及其表示矩阵，对称阵、二矩阵相合、正负惯标及正定矩阵的概念等。

本章的主要结果是：

1）任一对称阵总相合于对角形阵，进而相合于标准形，且标准形是唯一的。

2）任一二次型总可以经可逆变数变换化为平方和，进而化为标准形，且标准形是唯一的。

3）设 A 为 n 阶对称阵，下列说法是等价的，如

(a) A 是正定的；

(b) A 相合于单位阵 I_n；

(c) 对任一 $n\times 1$ 矩阵 $X\neq 0$，均有 $X'AX>0$；

(d) 存在可逆阵 P，使 $A=P'P$；

(e) A 的一切顺序主子式均大于 0。

本章的主要方法是：初等变换法。就是利用成套的初等变换化对称阵为标准形；还有配方法。

通过本章的学习，要求读者

1）正确理解与掌握本章的主要概念；

2）熟练而准确地将对称阵化成标准形，将二次型化成标准形，并会求出相合演化阵和可逆变数变换；

3）会判别一个阵是否正定，并会论证有关问题。

〔内容分析〕

§1　二次型　对称阵在相合之下的分类

本节由解析几何二次曲线、二次曲面用坐标变换化简问题，提出二次型用可逆变数变换化简问题。而二次型与对称阵建立了一对一的对应关系，于是二次型用可逆变数变换化简问题转化为对称阵在相合分类之下的标准形问题。这就点明了本章的中心问题。

通过这一节的学习，要求读者

1）正确理解和掌握二次型的定义、二次型表示矩阵的定义，二次型与对称阵之关系。

2）正确理解和掌握二矩阵相合的定义、相合关系的性质及利用相合关系给对称阵分类。

3）明确二次型用可逆变数变换化简与对称阵在相合之下的标准形之关系。

§2　对称阵在成套的初等变换下的化简

本节主要介绍一个结论，两个方法：一个结论是：任何对称阵都相合于对角形阵，进而任何对称阵都相合于特殊的对角形，即标

准形；任何实二次型 $f(x_1, \cdots, x_n)=X'AX$，都存在可逆变数变换 $X=PY$，使 $f(x_1, \cdots, x_n)=g(y_1, \cdots, y_n)$ 为平方和，进而存在可逆变数变换 $X=QZ$，使 $f(x_1, \cdots, x_n)=h(z_1, \cdots, z_n)$ 为标准形。两个方法：一是初等变换法；二是配方法。配方法的实质也是初等变换。

通过这一节的学习，要求读者

1）熟练地掌握用成套的初等变换化对称阵（实二次型）为对角形（平方和），进而化为标准形，并求出其相合演化阵（可逆变数变换）。

2）掌握用配方法化对称阵（实二次型）为对角形（平方和）及标准形，并求出其相合的演化阵（可逆变数变换）。

§3 惯性定律

本节主要讨论实、复对称阵在相合之下的标准形的唯一性问题。因为实对称阵在相合之下的标准形是对角线上为 ±1 或0的对角形阵，由于秩数是相合之下的不变量，所以非0元素的个数由阵唯一决定。剩下的就是正1的个数和负1的个数各是多少，惯性定律就是说正1的个数和负1的个数由阵所唯一确定。由于复对称阵的标准形是对角线上都为1的对角形，故由阵唯一确定。

通过这一节的学习，要求读者正确理解正负惯标的概念，明确实、复对称阵在相合之下标准形是唯一的。

§4 正定矩阵

本节主要讲两个问题：一是讲（半）正定和（半）负定矩阵的概念；二是讲（半）正定矩阵的性质和判别法：

1）A是正定的充分必要条件A相合于单位阵1_n；

2）A为（半）正定的充分必要条件是：任一非0 $n\times 1$ 矩阵

X，均有 $X'AX>0$（$X'AX\geqslant 0$）；

3）A为（半）正定的充分必要条件存在可逆阵P（存在 n阶方阵Q）使 $A=P'P$（$A=Q'Q$）；

4）A为（半）正定的充分必要条件是：A的顺序（一切）主子式大于（大于等于）零。

通过这一节的学习，要求读者

1）正确理解和掌握（半）正定、（半）负定矩阵的概念；

2）会判别一个对称阵是否（半）正定，会利用（半）正定矩阵的性质论证有关问题。

〔例题选解〕

例 1　若A为对称阵，则A^{-1}也是对称阵。

证明　因为 $(A^{-1})'=A'^{-1}=A^{-1}$，故 A^{-1} 为对称阵。

例 2　设 A、B 都是对称阵，那么 AB 一定是对称阵吗？

解　不一定，比如

$$A=\begin{pmatrix}1&2\\2&3\end{pmatrix},\qquad B=\begin{pmatrix}2&1\\1&1\end{pmatrix}$$

都是对称阵，但

$$AB=\begin{pmatrix}4&3\\7&5\end{pmatrix}$$

不是对称阵。

例 3　证明　n 阶对称阵A与B相合的充分必要条件是它们的标准形相同。

证明　必要性．由题设知，存在可逆阵P，使 $B=P'AP$．又设A的标准形为A_0，B 的标准形为 B_0，即存在可逆阵 Q，R，分别有

$$Q'AQ=A_0,\quad R'BR=B_0.$$

于是

$$B_0=R'BR=R'P'APR=R'P'(Q^{-1})'A_0Q^{-1}PR=S'A_0S,$$

其中 $S=Q^{-1}PR$。故 A_0 与 B_0 是相合的。因此$A_0=B_0$。

充分性。若 $A_0=B_0$，即 $Q'AQ=R'BR$，也就是

$$B=(R^{-1})'Q'AQR^{-1}=U'AU,$$

其中 $U=QR^{-1}$。故A与B相合。

例 4　证明：一个实二次型可分解为两个实系数一次齐次多项式之积的充分必要条件是它的秩等于 1 或秩等于 2 而符号差为 0 。

证明　必要性。设二次型 $f(x_1,\ \cdots,\ x_n)$ 可分解为

$$f(x_1,\ \cdots,\ x_n)=(a_1x_1+a_2x_2+\cdots+a_nx_n)(b_1x_1+b_2x_2+\cdots+b_nx_n) \tag{1}$$

其中 $a_i,\ b_j(i,j=1,\ \cdots,\ n)$ 均为实数。如果

$$\alpha=(a_1,\ a_2,\ \cdots,\ a_n),\ \beta=(b_1,\ b_2,\ \cdots,\ b_n)$$

线性相关，即存在实数 λ，使 $\alpha=\lambda\beta$。于是

$$f(x_1,\ \cdots,\ x_n)=\lambda(b_1x_1+b_2x_2+\cdots+b_nx_n)^2.$$

令可逆变数变换

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}=\begin{pmatrix} b_1 & b_2 & \cdots & b_n \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

代入（1）式，则有

$$f(x_1,\ \cdots,\ x_n)=g(y_1,\cdots,\ y_n)=\lambda y_1^2,$$

故 $f(x_1,\ \cdots,\ x_n)$ 的秩为 1 。

如果 $\alpha=(a_1,\ a_2,\ \cdots,a_n)$，$\beta=(b_1,\ b_2,\ \cdots,\ b_n)$ 不线性相关，不妨设

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\neq 0.$$

于是令可逆变数变换

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ b_1 & b_2 & b_3 & \cdots & b_n \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

代入（1）式，则有

$$f(x_1, \cdots, x_n) = g(y_1, \cdots, y_n) = y_1 y_2. \qquad (2)$$

再令可逆变数变换

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & 1 & & & \\ 1 & -1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

代入（2）式，则有

$$f(x_1, \cdots, x_n) = h(z_1, \cdots, z_n) = z_1^2 - z_2^2,$$

故 $f(x_1, \cdots, x_n)$ 的秩为 2，而符号差为 0。

充分性。如果 $f(x_1, \cdots, x_n)$ 的秩为 1，则其标准形为

$$f(x_1, \cdots, x_n) = \partial(y_1, \cdots, y_n) = \pm y_1^2.$$

令可逆变数变换

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \cdots a_n \\ & 1 & \\ & & \ddots \\ & & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad a_i \text{ 为实数}，a_1 \neq 0.$$

则有

$$f(x_1, \cdots, x_n) = \pm (a_1x_1 + a_2x_2 + \cdots + a_nx_n)^2.$$

如果 $f(x_1, \cdots, x_n)$ 的秩为 2，而符号差为 0，则 $f(x_1, \cdots, x_n)$ 的标准形为

$$f(x_1, \cdots, x_n) = g(y_1, \cdots, y_n) = y_1^2 - y_2^2$$
$$= (y_1 + y_2)(y_1 - y_2)。$$

令可逆变数变换

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ b_1 & b_2 & b_3 & \cdots & b_n \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \neq 0.$$

则有

$$\begin{aligned} f(x_1, \cdots, x_n) = [&(a_1+b_1)x_1+(a_2+b_2)x_2+\cdots \\ &+(a_n+b_n)x_n]\cdot[(a_1-b_1)x_1+(a_2-b_2)x_2 \\ &+\cdots+(a_n-b_n)x_n]. \end{aligned}$$

例 5　求实二次型

$$f(x_1, \cdots, x_{2n}) = x_1x_2 + x_3x_4 + \cdots + x_{2n-1}x_{2n}$$

的秩与符号差.

解　$f(x_1, \cdots, x_{2n})$的表示矩阵为

$$A = \begin{pmatrix} 0 & \frac{1}{2} & & & & & \\ \frac{1}{2} & 0 & & & & & \\ & & 0 & \frac{1}{2} & & & \\ & & \frac{1}{2} & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \frac{1}{2} \\ & & & & & \frac{1}{2} & 0 \end{pmatrix}.$$

由于

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

令

$$T=\begin{pmatrix}1&1&&&&&\\1&-1&&&&&\\&&1&1&&&\\&&1&-1&&&\\&&&&\ddots&&\\&&&&&1&1\\&&&&&1&-1\end{pmatrix},$$

则有

$$T'AT=\begin{pmatrix}1&0&&&&&\\0&-1&&&&&\\&&1&0&&&\\&&0&-1&&&\\&&&&\ddots&&\\&&&&&1&0\\&&&&&0&-1\end{pmatrix}$$

于是A 的秩为 $2n$，正惯标为 n，负惯标为 n，因此符号差为0。所以，二次型 $f(x_1, \cdots, x_{2n})$ 的秩为 $2n$，符号差为0。

例6　t 取何值时，二次型

$$f(x_1, x_2, x_3)=x_1^2+x_2^2+5x_3^2+2tx_1x_2-2x_1x_3+4x_2x_3$$

为正定的。

解　$f(x_1, x_2, x_3)$ 的表示矩阵为

$$A=\begin{pmatrix}1&t&-1\\t&1&2\\-1&2&5\end{pmatrix},$$

其各阶顺序主子式

$$|1|=1,\quad \begin{vmatrix}1&t\\t&1\end{vmatrix}=1-t^2=(1+t)(1-t),$$

$$|A|=-t(5t+4).$$

而由

$$(1+t)(1-t)>0,\ 得\ -1<t<1,$$

$$-t(5t+4)>0, \quad 得 -\frac{4}{5}<t<0。$$

故当 t 满足 $-\frac{4}{5}<t<0$ 时，A 为正定的。

例 7　设 A 为正定的。证明 A^{-1} 也是正定的。

证明　由 A 为正定的，必存在可逆阵 P，使

$$P'AP=I。$$

而

$$(P'AP)^{-1}=I。即 \ P^{-1}A^{-1}P^{-1\prime}=I。$$

说明 A^{-1} 相合于单位阵，故 A^{-1} 也是正定的。

例 8　若 A 为正定的，B 为非 0 方阵，则 $B'A^{-1}B$ 为半正定的。

证明　由上例知，A^{-1} 也是正定的。于是存在可逆阵 P，使 $A^{-1}=P'P$。这时一方面，有

$$(B'A^{-1}B)'=B'A^{-1}(B')'=B'A^{-1}B,$$

即 $B'A^{-1}B$ 为对称阵；另一方面

$$B'A^{-1}B=B'P'PB=(PB)'(PB)=Q'Q,$$

其中 $Q=PB$。故 $B'A^{-1}B$ 为半正定的。

第十章　方阵在相似之下的标准形

〔内容提要〕

本章讨论的中心问题是：方阵在相似分类之下的标准形和不变量问题．在二、三节给出了方阵相似于对角形阵的条件；五、六节解决了两种特殊类型的方阵——实对称阵、正交阵在正交相似之下的标准形问题；七、八节解决了一般方阵在相似之下的标准形和不变量问题．七节给出了具体求一个方阵的标准形的方法，八节给出了一般方阵的相似标准形和不变量的理论证明．第一节提出了本章讨论的主要课题，而第四节与二、三节一起为五、六节解决实对称阵、正交阵的正交相似标准形做了必要的准备．

本章的主要概念是：二方阵相似的概念；方阵的特征矩阵、特征多项式、最小多项式、特征根、特征向量、特征向量系的概念；正交矩阵、二向量正交、正交向量组、标准正交组，两个向量组正交的概念；不变因子、初等因子、行列式因子、可逆 λ-矩阵、若当型阵、有理型阵的概念．

本章的主要结果是

（一）方阵A相似于对角形阵的充分必要条件是A有完全特征向量系．若A的特征向量系为 α_1，α_2，…，α_n 时，令$P=(\alpha_1\ \alpha_2\ \cdots\ \alpha_n)$，那么

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},$$

其中 λ_1, …, λ_n 为A的全部特征根.

1）若 n 阶方阵A有 n 个不同特征根，则A有完全特征向量系. 在每个特征空间 $S_A(\lambda_i)$ 中取出一个特征向量 α_i，那么 $\alpha_1, \cdots, \alpha_n$ 就是A的完全特征向量系；

2）若每个特征空间 $S_A(\lambda_i)$ 的维数等于特征根 λ_i 的重数时，则A有完全特征向量系. 在每个特征空间 $S_A(\lambda_i)$ 中取出基底 u_{i1}, …, u_{il_i}, 那么

$$u_{11}, \cdots, u_{1l_1}, \cdots u_{t1}, \cdots, u_{tl_t}$$

就是A的完全特征向量系.

（二）任一对称阵A，总存在正交阵P，使

$$P^{-1}AP = P'AP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

其中 λ_1, …, λ_n 恰为A的全部特征根.

1）实对称阵A的特征根均为实数；

2）属于实对称阵A的不同特征根的特征向量是正交的；

由1）、2）立即得到

3）任一实对称阵A，必有完全特征向量系，并且有完全标准正交的特征向量系；

4）在A的每一特征空间 $S_A(\lambda_i)$ 中取出标准正交基底 $u_{i1}, \cdots, u_{il_i}$, 那么

$$u_{11}, \cdots, u_{1l_1}, \cdots, u_{t1}, \cdots, u_{tl_t}$$

就是A的完全标准正交的特征向量系. 若令

$$P = (u_{11} \cdots u_{1l_1} \cdots u_{t1} \cdots u_{tl_t}),$$

则有

$$P^{-1}AP = P'AP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

（三）任一正交阵A. 总存在正交阵T，使

$$T^{-1}AT = T'AT =$$

$$\begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & -1 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & -1 & & & & \\ & & & & & & \cos\theta_1 & -\sin\theta_1 & & \\ & & & & & & \sin\theta_1 & \cos\theta_1 & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \cos\theta_t & -\sin\theta_t \\ & & & & & & & & & \sin\theta_t & \cos\theta_t \end{pmatrix}.$$

其中±1为A的实特征根，而二阶块

$$\begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$$

由虚特征根的三角表示式所决定。

1）任一正交阵的特征根的模为1；

2）属于正交阵A的不同特征根的特征向量是正交的；

3）任一正交阵A，对于实特征根 λ_0，存在正交阵P，使

$$P^{-1}AP = \begin{pmatrix} \lambda_0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}.$$

其中 A_1 为正交阵，而P是以属于 λ_0 的标准化的特征向量做为第一列的正交阵；

4）任一正交阵A，对于虚特征根 $\lambda_0 = u + vi$ $(v \neq 0)$，存在正交阵P，使

$$P^{-1}AP = \begin{pmatrix} A_2 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}.$$

其中 A_2 是正交阵，而A的属于虚特征根 $\lambda_0 = u + vi$ 的特征向量分为实部和虚部系数，得到两个实 n 维向量 α，β，然后标准化为$\tilde{\alpha}$，$\tilde{\beta}$，以 $\tilde{\alpha}, \tilde{\beta}$ 为后两列的正交阵就是演化阵P。

（四）在任意数域上，任一n阶方阵A都相似于一个有理型阵N，可指定N的对角线上诸块的次序，使N是由A唯一决定；

在复数域上，任一n阶方阵A都相似于一个若当型阵J，如不计J的对角线上诸块的次序，J是由A唯一决定。

1）任一λ-矩阵$A(\lambda)$都相抵于唯一一个对角形λ-矩阵：

$$\begin{pmatrix} d_1(\lambda) & & & & \\ & \ddots & & & \\ & & d_r(\lambda) & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix},$$

其中$d_1(\lambda)$，…，$d_r(\lambda)$首系数为1且$d_1(\lambda)|d_2(\lambda)|\cdots|d_r(\lambda)$。

2）二λ-矩阵$A(\lambda)\longrightarrow B(\lambda)$充分必要条件是不变因子相同。

同秩二λ-矩阵$A(\lambda)\longrightarrow B(\lambda)$充分必要条件是初等因子组相同；

二λ-矩阵$A(\lambda)\longrightarrow B(\lambda)$充分必要条件是行列式因于相同；

二λ-矩阵$A(\lambda)\longrightarrow B(\lambda)$充分必要条件是有相同的标准形；

二λ-矩阵$A(\lambda)\longrightarrow B(\lambda)$充分必要条件：存在可逆$\lambda$-矩阵$P(\lambda)$，$Q(\lambda)$，使

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda).$$

3）二数字矩阵$A\sim B$的充分必要条件是：$\lambda I-A\longrightarrow\lambda I-B$。

4）A与其有理型阵N有相同的不变因子，于是$A\sim N$；

A与其若当型阵J有相同的初等因子组，于是$A\sim J$。

（五）n阶方阵在相似分类之下的不变量：

同秩的二数字矩阵$A\sim B$的充分必要条件有相同的不变因子；

同秩的二数字矩阵$A\sim B$的充分必要条件有相同的初等因子组；

同秩的二数字矩阵$A\sim B$的充分必要条件有相同的行列式因子。

（六）再补充数字矩阵A相似对角形阵的条件：

1）A相似于对角形阵的充分必要条件A的初等因子都是一次式；

2）A相似于对角形阵的充分必要条件A的最小多项式无重

根。

本章的主要方法是：

（一）求特征根、特征向量的方法。

1）解决一个阵是否相似于对角形时，一个重要的方法是求特征根、特征向量、特征向量系；

2）在解决实对称阵、正交阵的标准形时，是通过求特征根、特征向量、标准正交特征向量系解决的。

（二）Schmidt 正交化法。

Schmidt 正交化法是一个重要方法，在本章起了重要作用：

1）在解决实对称阵、正交阵的标准形问题上是不可缺少的；

2）构造标准正交组：一个线性无关向量组化为标准正交组；一个基底化为标准正交基底；一个含 $r(<n)$ 个向量的标准正交组 $\alpha_1, \cdots, \alpha_r$，可找一 α_{r+1}，使 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}$ 仍为标准正交组；

3）构造正交阵：一个可逆阵可以化为正交阵，一个秩为m的 $m\times n$ 矩阵 $(m<n)A$，可以先将行标准正交化，然后补上 $n-m$ 行使之成为正交阵。

（三）初等变换方法。

我们知道，初等变换方法是高等代数的重要方法，基本上是贯串高等代数始终。在这一章里，它也是一个重要方法。在解决一般方阵A的有理标准形、若当标准形时，一般说来先将其特征矩阵进行初等变换，求出不变因子，初等因子组，从而求出A的标准形。当然前面的求特征向量和构造正交阵都需要解方程组，而解方程组的重要方法是初等变换，可见初等变换方法在这一章也是贯串始终的。

通过这一章的学习，要求读者：

1）正确掌握本章的主要概念；

2）会判断一个方阵A是否相似于对角形，如果是，能够熟练而准确地求出所相似的对角形阵及演化矩阵；

3）对于实对称阵、正交阵会求它的标准形及演化矩阵；

4）会构造标准正交组、正交阵；

5）熟练而准确地求出方阵的有理标准形、若当标准形；

6）能够利用本章的有关概念和结果论证比较基本的问题。

〔内容分析〕

§1 方阵的相似及相似分类

本节首先从总结两种类型的矩阵分类和标准形问题——相抵、相合问题及线性变换的矩阵表示等提出了矩阵的相似分类和相似标准形问题；

其次进一步给出了二矩阵相似的定义和利用相似关系给出了矩阵的分类；

最后指明了本章讨论的主要课题：n 阶方阵在相似分类之下的标准形问题和不变量。

通过这一节的学习要求读者：

1）正确掌握矩阵相似的概念和利用相似关系给出矩阵分类；

2）明确本章讨论的主要课题。

§2 特征矩阵　特征向量

第二、三节主要讨论方阵A相似于对角形阵的条件，同时也为后来的讨论做必要的准备。

第二节　首先从n阶方阵A相似于对角形阵出发引进了n阶方阵A的特征矩阵，特征多项式、特征根、特征向量的概念，并由特征向量的定义立即得到一个重要事实：

非零的$n\times 1$矩阵 $\alpha=\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$

是A的属于特征根 λ 的特征向量必要而且只要

$$A\alpha=\lambda\alpha. \quad (1)$$

其次 给出了相似阵的一个性质：相似的两个矩阵的特征多项式、特征根均相同。说明特征多项式、特征根是相似不变量。这一性质以后经常用到。

最后，要着重指出的是：n 阶方阵A的特征根、特征向量的定义本身就给出了求法，并且已列举了三个例子具体示范了求一个阵的特征根、特征向量、特征空间。

通过本节的学习要求诸者：

1）正确掌握特征矩阵、特征多项式、特征根、特征向量的概念；

2）准确而又熟练的求一个阵的特征根,特征向量、特征空间；

3）恰当地运用有关概念和等式（1）会论证有关特征多项式、特征根、特征向量的命题。

这一节需要说明的：

1）一个方阵A的属于特征根 λ_i 的特征向量必是存在的，这是因为它是A的属于特征根 λ_i 的特征方程组的非零解，而这个方程组是齐次的且系数行列式 $|\lambda_i I-A|=0$，于是这个特征方程组必有非零解。

2）要充分认识等式（1），即$A\alpha=\lambda\alpha$ 的作用。它一方面可以做为方阵A的特征根、特征向量的定义，就是说：如果有一个非零的 $n\times 1$ 矩阵 β，满足

$$A\beta=\lambda\beta,$$

则称 λ 是A的特征根，而 β 是A的属于λ 的特征向量。很容易从这个定义出发，推出我们书上的定义，请读者自己把它推导出来。

另一方面，这个等式（1）在论证与特征根．特征向量有关的命题时要经常用到．请读者参看例题选解中的例2、例5．

3）本节命题的逆命题是不成立的，即A与B的特征多项式相同，但A与B不一定相似．比如

$$A=\begin{pmatrix}1 & \\ & 1\end{pmatrix},\quad B=\begin{pmatrix}1 & 1\\ & 1\end{pmatrix}$$

的特征多项式都是$(\lambda-1)^2$，但A与B不相似．

4）求方阵A的特征向量的步骤如下：

（a）求A的特征多项式；

（b）求A的特征根；

（c）分别求属于各特征根的特征向量．这个步骤首先需要对于任一特征根写出其特征方程组，然后解之即求出了特征向量．

5）方阵A的特征根、特征向量与所讨论的数域有关．因为，矩阵A的特征根是A的特征多项式的根，而多项式的求根问题与数域是有关的，从而特征向量也就与数域有关了．比如：设三阶方阵

$$A=\begin{pmatrix}5 & 6 & -3\\ -1 & 0 & 1\\ 1 & 2 & -1\end{pmatrix},$$

求A的特征根、特征空间、特征向量．

解　A的特征多项式为

$$f(\lambda)=|\lambda I-A|=\begin{vmatrix}\lambda-5 & -6 & 3\\ 1 & \lambda & -1\\ -1 & -2 & \lambda+1\end{vmatrix}$$

$$=(\lambda-2)(\lambda^2-2\lambda-2)．$$

（a）如在有理数域中讨论，那么A只有一个特征根2．经计算，得

$S_A(2)=\{k(2,\ -1,\ 0)\mid k$ 为任意有理数$\}$；

A的属于2的特征向量为

$(2k,\ -k,\ 0)$，k 为非0有理数．

（b）如在实数域中讨论，那么A除有特征根2外还有两个特征根 $1+\sqrt{3}$，$1-\sqrt{3}$．于是除 $S_A(2)$ 外还有 $S_A(1+\sqrt{3})$，$S_A(1-\sqrt{3})$．经计算，得

$S_A(1+\sqrt{3})=\{k(6+3\sqrt{3},\ -2-\sqrt{3},\ 1)\mid k$为任意实数$\}$；

$S_A(1-\sqrt{3})=\{k(6-3\sqrt{3},\ -2+\sqrt{3},\ 1)\mid k$为任意实数$\}$；

A的属于$1+\sqrt{3}$，$1-\sqrt{3}$的特征向量分别为

$((6+3\sqrt{3})k,\ (-2-\sqrt{3})k, k)$，　$((6-3\sqrt{3})k,\ (-2+\sqrt{3})k, k)$，其中 k 为非0实数．

请读者列举一个阵，它在实数域中没有特征根，没有特征向量，而在复数域中有特征根、特征向量．

§3　特征向量系

本节在第二节讨论的基础上，将要明确指出一个方阵A相似于对角形阵的条件．为此

首先分析了A相似于对角形阵的条件，指出A相似于对角形阵的充分必要条件看其是否存在由A的n 个线性无关的特征向量做成的演化矩阵．

其次为了解决一个方阵A是否存在 n个线性无关的特征向量做成的演化矩阵P，我们证明了三个命题．这三个命题说明了A的特征向量系所含向量的个数是A的所有特征向量中线性无关的向量的最大个数．这个最大个数为 n时（也就是A的各特征空间的维数均等于其特征根的重数），这个特征向量系是完全的．这时以这个完全的特征向量系的各个向量为列做成的阵P，即是所要求的演化矩阵．于是得到了n 阶阵A相似于对角形的充分必要条件是：A有完全的特征向量系．

本节最后，为了讨论相似标准形的需要，介绍了哈米顿-凯莱(Hamilton-Cayley) 定理、最小多项式及其性质与求法．

通过这一节的学习要求读者：

1）正确掌握特征向量系的概念，明确一个n阶方阵相似于对角形阵的充分必要条件，会求特征向量系、演化矩阵及阵A所相似的那个对角形；

2）能够掌握最小多项式的概念、性质和求法；

3）会运用特征向量的性质及最小多项式论证有关问题.

这一节需要说明的：

1）一个方阵A的特征向量系不是唯一的. 这是因为，从命题3的证明过程已看到，特征向量系是由A的各个特征空间的基底构成的，而特征空间的基底的取法是不唯一的. 又由于各特征空间的维数是确定的，即基底所含向量的个数是确定的，因此A的特征向量系所含向量的个数是由A唯一决定. 于是，如果A有一个完全特征向量系，那么A的所有特征向量系都是完全的.

2）一个方阵A，如果相似于对角形阵，那么它的演化矩阵P是不唯一的. 这是因为A的完全特征向量系不唯一.

又因为与A相似的对角形阵，对角线上的元素是A的全部征特根. 所以，与A相似的对角形阵，如不计对角线上诸元素的次序是由A唯一决定.

3）求方阵A的特征向量系、相似的对角形阵和演化矩阵的步骤：

（i）求A的所有特征根，以特征根作为对角线上的元素的对角形阵就是与A相似的对角形阵；

（ii）求出各特征空间 $S_A(\lambda_i)$ 的基底，把这些基底放在一起就构成A的特征向量系；

（iii）以特征向量系的各个特征向量做列所构成的阵就是所求的演化矩阵P. 值得注意的是：演化矩阵P的各列的顺序一定要和A相似的那个对角形阵D的对角线上的元素的顺序相一致，即P的第i列是D的i行i列元素的特征向量，这时才有

$$P^{-1}AP=D.$$

4）由于A的特征根、特征向量与所讨论的数域有关，因此阵

A的特征向量系也与数域有关，进而A是否有完全特征向量系也与数域有关，故A能否与对角形阵相似也与数域有关。

§4 正交矩阵

本节为了给五、六节讨论对称阵、正交阵在正交相似之下的标准形做准备，这一节讨论正交阵的一般性质和构造正交阵的方法。

本节前半段是给出正交阵的定义和正交阵的性质。

命题2的第二条："正交阵A的特征根的模等于1"，在讨论正交阵的标准形时是必不可少的；命题3从阵的元素角度进一步揭示了正交阵的特性：每行（列）元素平方和等于1，任二不同行（列）对应元素乘积之和为0。（当后半段给出标准正交组定义后，正交阵的定义又可叙述为：行（列）向量为标准正交组）。这一特性又可做为正交阵的定义，同时也为后半段构造正交矩阵奠定了基础。

本节后半段，首先给出了两个n维向量正交的定义，进而给出了标准向量组、正交向量组、标准正交组的定义及一个向量和一组向量正交、组与组正交的概念。

其次由定理1给出了从一组线性无关的向量组构造标准正交组的方法，这个事实有以下的等价说法：

1）对给定的秩为r的$r\times n$矩阵$A=(a_{ij})$能够找到r阶下三角形阵P，使$B=PA$的行向量组为标准正交组；

2）对于给定的秩为r的$n\times r$矩阵$A=(a_{ij})$，存在r阶上三角形阵Q，使$B=AQ$的列向量组为标准正交组；

3）对任一可逆阵A，存在下（上）三角形阵$P(Q)$，使$T_1=PA$（$T_2=AQ$）为正交阵。就是说，从一个可逆阵可得到一个正交阵，这是构成正交阵的一种方法。

第三　由命题4及定理1给出了n维空间V存在由标准正交组构成的基底，这个标准正交基底可由空间任一基底，按着定理1的

办法构造出来。

第四，由命题5和定理1给出了在一个标准正交组的基础上添加向量，使其仍是标准正交组：

一是利用命题5；二是利用定理1，即设标准正交向量组α_1，α_2，… α_r，再找一向量β，使

$$\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_r,\ \beta$$

线性无关。然后按定理1使其标准正交化。

这样一来，对于给定的标准正交组总可以扩充为n维空间的标准正交基底。换一种说法，一个给定的 $r\times n$（$r<n$）的长方矩阵T_0，若其行（列）向量组是标准正交的，那么总可以找到$(n-r)\times n$ 矩阵 T_1，使

$$T=\begin{pmatrix}T_0\\T_1\end{pmatrix}$$

成为n阶正交阵。这又获得了从一个秩等于行（列）数的长方矩阵构成一个正交矩阵的方法：先把长方矩阵的行（列）标准正交化使其成为 T_0，然后再求 T_1，使

$$\begin{pmatrix}T_0\\T_1\end{pmatrix}$$

为正交阵。

通过本节的学习，要求读者：

1）正确掌握正交阵的概念及其性质；

2）正确掌握两个向量正交，正交向量组，标准正交组，二向量组正交等概念及正交向量组的性质；

3）熟练地掌握 Schmidt 正交化法，会构造标准正交向量组、标准正交基底、正交阵；

4）能够运用正交阵及其性质，正交向量组及其性质以及Schmidt 正交化论证有关问题。

本节需要说明的：

1）据标准正交向量组的定义知，α_1，…，α_r 为标准正交组必

要而且只要

$$\alpha_i \alpha_j' = \begin{cases} 1, & i=j \\ 0, & i \neq j; \end{cases}$$

而 α_1 …, α_n 为标准正交基底必要而且只要

$$\alpha_i \alpha_j' = \begin{cases} 1, & i=j \\ 0, & i \neq j. \end{cases}$$

这两个事实，经常用来论证或判断一组向量是否是标准正交组（基底）。

2）两个向量组正交，并不意味着每个组都是正交的。比如向量组

$$\alpha_1 = (1,0,0),\ \alpha_2 = (1,1,0),\ \alpha_3 = (2,2,0)$$

与

$$\beta_1 = (0,0,1),\ \beta_2 = (0,0,2)$$

正交，但向量组 α_1，α_2，α_3 不是正交组；β_1，β_2 也不是正交组。

3）正交阵必须是实方阵，且是可逆阵，命题2的第一条，“正交阵的行列式必为±1”，其逆是不成立的。比如

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

的行列式是1，但A不是正交阵。

4）定理1的证明，可对向量的个数用归纳法，而使证明过程会显得精练些，具体过程请读者给出。

5）在定理2中，使$\begin{pmatrix} T_0 \\ T_1 \end{pmatrix}$成为正交阵的 T_1 不是由 T_0 所唯一决定的。换句话说，由已知 T_0 可以构造出不止一个正交阵。这是因为方程组

$$T_0 X = 0$$

的解空间的标准正交基底不止一个，并且有无穷多个。故可找无穷多个T_1，使$\begin{pmatrix} T_0 \\ T_1 \end{pmatrix}$为正交阵。

§5 实对称阵在正交相合之下的标准形

从本节开始，进入这一章的中心课题，即方阵在相似之下的标准形问题. 本节是研究特殊的矩阵——实对称阵在正交相合（相似）之下的标准形. 要得到的主要结果是：任一实对称阵A，总存在正交阵 P，使 $P'AP=P^{-1}AP$ 为对角形阵. 与之等价的说法就是：任一实对称阵，必有完全标准正交的特征向量系，且每个特征向量都是实数组. 为此

首先通过命题1，确定实对称阵A的特征根均为实数. 这就保证了阵A的属于每个特征根的特征方程组的系数均为实数，故阵A的特征向量均为实数组. 这样，若阵A有完全的标准正交特征向量系，则以每个向量作列构成的阵才是正交阵.

其次由命题2保证了，当在A的各个特征空间均取标准正交基底，把这些基底放在一起就是标准正交特征向量系.

第三由定理1知，实对称阵A的特征向量系是完全的，因此立即得到了定理2，从而得到了本节所要的结果.

第四最后的推论是重要的，推论1指明了实对称阵所相似的对角形不是别的，它的对角线上的元素恰是A的全部特征根；推论2说明：一个实对称阵是否是正定的、半正定的可以看其特征根是否均为正，还是均为非负而定，这里又获得了判断实对称阵正（负）定、半正（负）定的一个方法；推论3给出了化实二次型为平方和的一个方法，即正交变换法.

通过本节学习，要求读者：

1）对于给定的实对称阵A，能熟练而准确地求出A的完全标准正交的特征向量系，进而求出正交阵P，使 $P'AP=P^{-1}AP$ 为对角形阵；

2）会用正交变换法化实二次型为平方和，并求出所用的可逆变数变换；

3）能够运用实对称阵的特征根、特征向量的性质论证有关问题。

这一节需要说明的问题：

1）对给定的实对称阵A，求正交阵P，使 $P'AP=P^{-1}AP$ 为对角形的步骤：

（i）求A的特征根，于是A就相似于这些特征根作为对角线上的元素的对角形阵；

（ii）求各特征空间的基底，也就是求A的各个特征方程组的基础解系；

（iii）将各特征空间的基底标准正交化，然后放在一起就构成了完全标准正交特征向量系；

（iv）把特征向量系中各向量作列，便构成了正交阵，就是要求的演化阵P。

2）我们知道，对实对称阵A，总存在正交阵P，使 $P'AP=P^{-1}AP$ 为对角形，称为A 的标准形。那么正交阵P是否唯一，标形准是否唯一？

首先，演化阵P不是唯一的。因为P是由A的完全标准正交特征向量系所决定的，而A的完全标准正交特征向量系不唯一；

其次，A的标准形是由A的特征根所决定，于是若不计对角线上各元素的顺序，A的标准形是由A唯一确定。

值得注意的是，A的标准形的对角线上元素的顺序与正交阵P各列的顺序有关。

§6 正交矩阵在正交相似之下的标准形

本节是讨论正交阵在正交相似之下的标准形。也就是，对任一正交阵A，如何选取正交阵T，可使 $T^{-1}AT=T'AT$ 得到怎样简单的形式？为此，首先证明了命题1、2，然后得到了标准形定理。至于命题1、2起了什么作用，书中已说得十分清楚，这里不

再重述，需要说明的：

1）命题2的证明比较长，读者不易抓住证明的思路，我们在这略述一下。要证明的是，当正交阵 A 有一虚根时，寻找演化阵 P，把 A 演化为分块阵：

$$\begin{pmatrix} A_2 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}. \tag{1}$$

其主要思路是

i）如何选取演化阵 P；

ii）选得 P 后，验证 P 确实能把 A 演化成（1）型。

先说选取 P。因特征根 λ_0 为虚根，故属于 λ_0 的特征向量 γ 是复 n 数组，把向量 γ 分成：

$$\gamma = \alpha + \beta i,$$

其中

$$\alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \qquad \beta = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

为两个实 $n\times 1$ 矩阵。下面，第一步证 α，β 是线性无关的；第二步把 α，β 标准正交化为 $\tilde{a}$，β^*；第三步选取以 $\tilde{a}$，β^* 为最后两列的正交阵 P；

其次验证

$$P^{-1}AP = P'AP = \begin{pmatrix} A_2 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}.$$

第一步算出

$$P^{-1}AP = P'AP = \begin{pmatrix} A_2 & 0 \\ A_1 & C \end{pmatrix},$$

其中 $|C|=1$；第二步利用 $P^{-1}AP$为正交阵的性质，算出 $A_3=0$；第三步由 $|C|=1$，可知

$$C=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix},$$

故命题 2 得证。

2）求正交阵A的标准形及演化矩阵P的步骤：

（a）求A的特征根。若有实特征根必为±1，虚特征根成对：

$$1,\ \cdots,\ 1,\ -1,\ \cdots,\ -1,\ u_1+v_1i,\ u_1-v_1i,\ \cdots,$$
$$u_r+v_ri,\ u_r-v_ri.$$

（b）求A的标准形。将每个虚根 u_1-v_1i，…，u_r-v_ri 表成三角形式：

$$u_1-v_1i=\cos\theta_1+i\sin\theta_1,\quad\cdots,\quad u_r-v_ri=\cos\theta_r+i\sin\theta_r.$$

于是A的标准形为

$$\begin{pmatrix}1&&&&&&&&&&\\&\ddots&&&&&&&&&\\&&1&&&&&&&&\\&&&-1&&&&&&&\\&&&&\ddots&&&&&&\\&&&&&-1&&&&&\\&&&&&&\cos\theta_1&-\sin\theta_1&&&\\&&&&&&\sin\theta_1&\cos\theta_1&&&\\&&&&&&&&\ddots&&\\&&&&&&&&&\cos\theta_r&-\sin\theta_r\\&&&&&&&&&\sin\theta_r&\cos\theta_r\end{pmatrix}.$$

（c）求演化矩阵P。

首先，求A的属于 1 的特征方程组的基础解系，然后标准正交化做为P的前几列；

其次，求A的属于 -1 的特征方程组的基础解系，然后标准正交化做为P的接下去的几列；

第三，求A的属于 u_1-v_1i 的特征向量 γ_1，将γ_1分解为实部、虚部：$\gamma_1=\alpha_1+\beta_1i$，然后将 α_1，β_1 标准正交化后做P 的接下去的列，这样下去，求A的属于 u_r-v_ri 的特征向量 γ_r，将γ_r分解为：

$\gamma_r=\alpha_r+\beta_r i$，把 α_r，β_r 标准正交化后做为 P 的最后两列。

由于正交阵 A 的属于不同特征根的特征向量是正交的（练习六第 3 题），于是 P 为正交阵且使 A 化为标准形的演化矩阵。

3）正交阵 A 标准形的唯一性问题，演化阵 P 的唯一性问题。

正交阵 A 的标准形对角线上各块，由 A 所唯一决定，但对角线上各块的次序可以不同，故如不计对角线上各块的次序，正交阵 A 的标准形是唯一的。至于使 A 为标准形的演化阵 P，是由 A 的特征向量做成的，P 不是唯一的。

4）二、三阶正交阵的几何意义。

从正文十章§6例 1 可知，二阶正交阵的标准形就是下列两种形式：

$$\begin{pmatrix}1 & \\ & -1\end{pmatrix},\quad \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

今将 $(x,\ y)$ 看成平面上一点 P 的坐标，于是变换

$$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

是平面上绕原点旋转 θ 角，把点 $P(x,\ y)$ 变为 $Q(x',\ y')$ 的变换。而变换

$$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}1 & \\ & -1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

是关于 x 轴的反射，把点 $P(x,\ y)$ 变为 $Q(x',\ y')$。

三阶正交阵的标准形式为下列六种：

$$(1)\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix},\qquad (2)\begin{pmatrix}1 & & \\ & 1 & \\ & & -1\end{pmatrix},$$

$$(3)\begin{pmatrix}1 & & \\ & -1 & \\ & & -1\end{pmatrix},\qquad (4)\begin{pmatrix}-1 & & \\ & -1 & \\ & & -1\end{pmatrix},$$

(5) $\begin{pmatrix} 1 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}$, (6) $\begin{pmatrix} -1 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix}$.

（1）是恒等变换，即使空间的每个点都不动；

（2）是以 xoy 平面为对称平面的反射；

（3）是以 x 轴为对称轴的反射；

（4）是以原点为对称中心的反射；

（5）是以 x 轴为旋转轴，旋转 θ 角的变换；

（6）是以 x 轴为旋转轴，旋转 θ 角，接着再以 yoz平面为对称平面的反射.

但是，前四种可以看成后二种的特殊情况. 因为在（5）中，令 $\theta=0$ 即得（1），令 $\theta=\pi$ 即得（3）；在(6)中，令$\theta=0$，即得（2）（只相差对角线上的次序），令 $\theta=\pi$，即得（4）.

总起来说，三阶正交阵或者是以一直线为轴的旋转，或者是以一直线为轴的旋转再伴随着一个以平面为对称平面的反射.

这一节要求读者：能够掌握正交矩阵标准形的形式，会求其标准形和演化矩阵.

§7 有理标准形 若当标准形

上边两节是在实数域上讨论特殊的n 阶方阵在相似之下的标准形问题，所用的工具主要是特征根、特征向量. 这一节和下一节在一般数域和复数域上讨论方阵在相似之下的标准形问题，所用的工具与上两节有明显的不同，将要用不变因子、初等因子来解决标准形问题. 这一节主要介绍两个概念，即不变因子、初等因子；两个方法，即求相似标准形的方法：一是求有理标准形的方法，二是求若当标准形的方法. 这两种方法正文写得比较详细，这里不多说了.

通过这一节的学习，要求读者正确理解不变因子、初等因子的

概念；掌握求有理标准形和若当标准形的方法.

这一节需要说明的：

1）不管求有理标准形，还是求若当标准形，都需要求A的不变因子，也就是必须把 $\lambda I-A$ 化为特殊的对角形，所用的方法是初等变换. 那么用初等变换把 λ-矩阵 $\lambda I-A$ 化为适合要求的对角形，总的思路和方向是什么呢？和数阵用初等变换化对角形的思路方向一致. 那就是，首先在阵的左上角即(1,1) 位置找一个次数最低的非零多项式且能整除它所在的行、列其余元素，然后利用消法变换把所在的行、列其余元素化为 0；其次按照上述办法重复进行直到化成对角形为止；第三检查对角线上元素是否依次一个整除一个，如不然再进行消法变换直到满足这个条件为止. 如有首系数不是 1 的，可用倍法变换变为 1. 于是就是所求的形式了.

2）我们知道，给定方阵A，求出的若当阵 B 是上三角形阵. 而上三角形阵 B 的特征根恰是对角线上的元素，这些元素是A的初等因子组决定的，每个初等因子的一次式中那个根就是 B 的特征根. 由 $A\sim B$，所以 B 的特征根恰是A的特征根. 故 B 的对角线上的元素恰是A的全部特征根，也就是A的初等因子中所有一次式的那个根都是A的特征根，如果重复次数也计算在内，恰好是A的全部特征根.

§8 λ-矩阵在初等变换下的化简　相似定理

本节是本章最后一节，主要解决相似标准形的理论问题：

第一　证明任一方阵A的特征矩阵 $\lambda I-A$可经若干次初等变换化为

$$B(\lambda)=\begin{pmatrix} d_1(\lambda) & & & \\ & d_2(\lambda) & & \\ & & \ddots & \\ & & & d_n(\lambda) \end{pmatrix} \tag{1}$$

其中 $d_i(\lambda)$ 都是首系数为 1 且 $d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_n(\lambda)$. 并且进一步

证明（1）形式是唯一的；

第二　证明任一方阵A都相似于它的有理标准形，也相似于它的若当标准形。

本节从解决上述二个问题出发，做了一系列的准备工作。接着首先证明了任一λ-阵 $A(\lambda)$ 都相抵于一个对角形 λ-阵。

$$\begin{pmatrix} d_1(\lambda) & & & & & \\ & \ddots & & & & \\ & & d_r(\lambda) & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}$$

其中 $d_i(\lambda)$ 为首系数是1且 $d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_r(\lambda)$。特别是由此得到两个重要推论，即推论1和推论2，尤其推论2在论证λ-阵相抵的有关问题经常用到。

为了解决相抵之下标准形的唯一性，引入了行列式因子的概念并证明了行列式因子在相抵之下是不变量，从而解决了标准形的唯一性问题。于是得到

1）行列式因子与不变因子之关系：

$$D_1(\lambda)=d_1(\lambda),\quad D_2(\lambda)=d_1(\lambda)d_2(\lambda),\quad\cdots$$

$$D_r(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_r(\lambda);\quad d_1(\lambda)=D_1(\lambda),$$

$$d_2(\lambda)=\frac{D_2(\lambda)}{D_1(\lambda)},\quad\cdots,\quad d_r(\lambda)=\frac{D_r(\lambda)}{D_{r-1}(\lambda)}.$$

2）不变因子在初等变换之下不变，就是说不变因子是相抵之下的不变量。

3）据初等因子与不变因子之关系，初等因子在初等变换之下不变，即初等因子也是相抵之下的不变量。

总起来，得到了

同秩 λ-阵$A(\lambda)\longrightarrow B(\lambda)$ 充分必要条件是有相同的不变因子；

同秩 λ-阵 $A(\lambda)\longrightarrow B(\lambda)$ 充分必要条件是有相同的初等因子组；

同秩 λ- 阵 $A(\lambda) \longrightarrow B(\lambda)$ 充分必要条件是有相同的行列式因子；

同秩λ-阵 $A(\lambda) \longrightarrow B(\lambda)$ 充分必要条件是有相同的标准形；

同秩λ-阵 $A(\lambda) \longrightarrow B(\lambda)$ 充分必要条件：存在可逆 λ-阵$P(\lambda)$，$Q(\lambda)$，使

$$B(\lambda) = P(\lambda) A(\lambda) Q(\lambda).$$

其次有了定理 4，可以总结求不变因子和初等因子组的方法如下：

1）通过初等变换，将λ-阵化为标准形，求不变因子，再求初等因子组；

2）通过初等变换，将λ-阵化为对角形，先求对角线上各多项式的标准分解式，从而求出初等因子组，再求不变因子；

3）先求λ-阵的行列式因子，然后求不变因子，再求初等因子组。

一般说来，2）是比较方便些，因为化对角形还是易于做到的。但λ-阵比较特殊，行列式因子好求，那当然用3）为宜。比如象 n阶若当块

$$J_i = \begin{pmatrix} \rho_i & 1 & & \\ & \rho_i & 1 & \\ & & \ddots & \ddots \\ & & & 1 \\ & & & \rho_i \end{pmatrix}$$

的特征矩阵

$$\lambda I - J_i = \begin{pmatrix} \lambda-\rho_i & -1 & & \\ & \lambda-\rho_i & -1 & \\ & & \ddots & \ddots \\ & & & -1 \\ & & & \lambda-\rho_i \end{pmatrix}$$

有一 $n-1$ 阶子式

$$\begin{vmatrix} -1 & & & \\ \lambda-\rho_i & -1 & & \\ & \ddots & \ddots & \\ & & & -1 \\ & & & \lambda-\rho_i \end{vmatrix} = (-1)^{n+1}.$$

于是 $\lambda I - J_i$ 的 $n-1$ 阶的行列式因子 $D_{n-1}(\lambda)=1$，故

$$D_1(\lambda)=1,\ \cdots,\ D_{n-1}(\lambda)=1,\quad D_n(\lambda)=(\lambda-\rho_i)^n.$$

因而不变因子为

$$d_1(\lambda)=1,\ \cdots,\ d_{n-1}(\lambda)=1,\quad d_n(\lambda)=(\lambda-\rho_i)^n.$$

其初等因子组为 $(\lambda-\rho_i)^n$.

当给定λ-阵开始不好求行列式因子，可先进行初等变换，化到一定时候行列式因子易求了，这时也可用3），即2）与3）结合起来用。

第三　为了解决本节的第二个问题，先通过证明命题3、命题4解决了方阵A与其有理型阵有相同的不变因子，A与其若当型阵有相同的初等因子组。从而得到了A的特征矩阵与其有理型阵、若当型阵的特征矩阵相抵。接着通过命题5、6的准备证明了二数字矩阵相似的充分必要条件其特征矩阵相抵，从而得到了

$$A \sim N,\ A \sim J.$$

这里N为A的有理标准形，J为A的若当标准形。

第四　通过解决本节的两个问题，得到了

两个数字矩阵 $A\sim B$的充分必要条件不变因子相同；$A\sim B$ 的充分必要条件初等因子组相同；$A\sim B$ 的充分必要条件行列式因子相同；$A\sim B$ 的充分必要条件 $\lambda I-A$ 与 $\lambda I-B$ 的标准形相同；$A\sim B$ 的充分必要条件：存在可逆λ-阵 $P(\lambda)$，$Q(\lambda)$，使

$$\lambda I - B = P(\lambda)(\lambda I - A)Q(\lambda).$$

上述这些条件是判别两个数字矩阵是否相似的条件。当两个矩阵是具体给出时，要判别是否相似，一般采用前四个条件的某一个；当两个阵都不是具体给出，要证明两个阵相似一般用最后这个条件，当然也可能用前几个条件。

第五　利用初等因子、不变因子及最小多项式，得出了判别方阵A与对角形相似的条件.

1） A与对角形阵相似的充分必要条件为A的初等因子都是一次的；

2） A与对角形阵相似的充分必要条件是A的最小多项式无重根.

通过这一节的学习，要求读者

1） 正确理解和掌握不变因子、初等因子、行列式因子的概念及其相互关系以及可逆λ-阵与初等λ-阵之关系；

2） 掌握两个λ-阵相抵的条件，且会判断两个λ-阵是否相抵；

3） 掌握两个数字矩阵相似的条件，且会判断两个数字矩阵是否相似；

4） 能够运用相抵的条件、相似的条件以及若当标准形论证有关问题.

这一节还需说明的：

1） 数字矩阵A的有理标准形，如不计对角线上诸块的次序是由A唯一决定的. 这因为不变因子由A唯一决定，每个不变因子的伴侣阵由那个不变因子唯一决定，剩下只是有理型阵对角线上诸伴侣阵的次序先后有所不同而已，据命题3和定理5，各伴侣阵任一次序所构成的有理型阵都相似于A.

2） 数字矩阵的若当标准形，如果不计对角线上诸块的次序，是由A所唯一决定的. 这是因为A的初等因子组是由A唯一决定，而每个初等因子唯一决定一个若当块，剩下的只是各个若当块在若当型阵的对角线上先后次序不同而已，据命题4与定理5，各个若当块任一次序所构成的若当型阵都相似于A.

3） n 阶数字矩阵A的特征矩阵的行列式恒不为0，即$|\lambda I_n - A| \neq 0$. 于是A的行列式因子 $D_n(\lambda) = |\lambda I_n - A| \neq 0$.

〔例题选解〕

例1　若 $B_1=P^{-1}A_1P$，$B_2=P^{-1}A_2P$，则 $A_1+A_2\sim B_1+B_2$，$A_1A_2\sim B_1B_2$，$kA_1\sim kB_1$；当 $f(x)$ 为任一多项式，且 $B=P^{-1}AP$，那么 $f(B)=P^{-1}f(A)P$。其中 k 为常数。

证明　因为 $B_1+B_2=P^{-1}A_1P+P^{-1}A_2P=P^{-1}(A_1+A_2)P$，所以 $A_1+A_2\sim B_1+B_2$。同理可证 $A_1A_2\sim B_1B_2$，$kA_1\sim kB_1$。

又设 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$，那么

$$\begin{aligned}f(B)&=a_0B^n+a_1B^{n-1}+\cdots+a_{n-1}B+a_nI\\&=a_0(P^{-1}AP)^n+a_1(P^{-1}AP)^{n-1}+\cdots+a_{n-1}P^{-1}AP+a_nI\\&=a_0P^{-1}A^nP+a_1P^{-1}A^{n-1}P+\cdots+a_{n-1}P^{-1}AP+a_nP^{-1}IP\\&=P^{-1}(a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI)P\\&=P^{-1}f(A)P.\end{aligned}$$

例2　如果 n 阶阵A适合 $A^2=A$，则称A为幂等阵。证明幂等阵的特征根只有1或0。

证明　假设 λ 是A的特征根，α 是A的属于 λ 的特征向量，那么必有 $A\alpha=\lambda\alpha$。于是

$$A\alpha=A^2\alpha=A(A\alpha)=A(\lambda\alpha)=\lambda(A\alpha)=\lambda^2\alpha.$$

因此，有

$$\lambda\alpha=\lambda^2\alpha \text{ 或 } (\lambda-\lambda^2)\alpha=0.$$

由于 $\alpha\neq0$，所以必有 $\lambda-\lambda^2=0$，$\lambda(\lambda-1)=0$，故 $\lambda=0$，或 $\lambda=1$。

例3　两个矩阵有完全相同的特征根时，那么它们的特征向量也相同吗？

解　不一定，比如

$$A=\begin{pmatrix}0&-2\\1&3\end{pmatrix},\qquad B=\begin{pmatrix}-2&-2\\6&5\end{pmatrix}$$

的特征根均为1，2，但A的属于1、2的特征向量分别为

$$k\alpha_1=k(-2,\ 1)=(-2k,k)$$

$$k\alpha_2 = k(-1,\ 1) = (-k,\ k),\quad k \text{ 为非 0 常数.}$$

而 B 的属于 1、2 的特征向量分别是

$$k\beta_1 = k(-2,\ 3) = (-2k,\ 3k)$$

$$k\beta_2 = k(-1,\ 2) = (-k,\ 2k),\quad k \text{为非 0 常数.}$$

可见 A 与 B 的特征向量都不同，故 A, B 没有相同的特征向量.

例 4　设 n 阶方阵 A 有 n 个不同的特征根 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 求 $\lambda_i I - A$ 的秩，$i = 1, 2, \cdots, n$.

解法一　由题设知，每个 λ_i 都是单根，故 $S_A(\lambda_i)$ 的维数为 1，所以对应于 λ_i 的特征方程组

$$(\lambda_i I - A)X = 0$$

的基础解系所含向量的个数为 1，于是 $\lambda_i I - A$ 的秩为 $n-1$.

解法二　由题设知，存在可逆阵 P，使

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

于是，一方面

$$\lambda_i I - P^{-1}AP = \begin{pmatrix} \lambda_i - \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_i - \lambda_{i-1} & & & \\ & & & 0 & & \\ & & & & \lambda_i - \lambda_{i+1} & \\ & & & & & \ddots \\ & & & & & & \lambda_i - \lambda_n \end{pmatrix}$$

的秩为 $n-1$. 另一方面，由

$$\lambda_i I - P^{-1}AP = P^{-1}(\lambda_i I - A)P,$$

说明 $\lambda_i I - A$ 的秩为 $n-1$.

例 5　设 α、β 分别是 A 的属于不同特征根 $\lambda_1, \lambda_2 (\lambda_1 \neq \lambda_2)$ 的特征向量. 证明 $\alpha + \beta$ 不是 A 的特征向量.

证明　反证法. 假设 $\alpha + \beta$ 是 A 的特征向量，那么必有

$$A(\alpha+\beta) = \lambda(\alpha+\beta),\ \text{或}\ A\alpha + A\beta = \lambda\alpha + \lambda\beta.$$

又 $A\alpha = \lambda_1\alpha$, $A\beta = \lambda_2\beta$. 于是 $A\alpha + A\beta = \lambda_1\alpha + \lambda_2\beta$. 故有

$$\lambda\alpha+\lambda\beta=\lambda_1\alpha+\lambda_2\beta,\ 或\ (\lambda-\lambda_1)\alpha+(\lambda-\lambda_2)\beta=0. \qquad (*)$$

由于 α, β 是A的属于不同特征根的特征向量，因而是线性无关的。因此由(*)，必有

$$\lambda-\lambda_1=0,\ \lambda-\lambda_2=0\ 或\ \lambda_1=\lambda_2\ 与\ \lambda_1\neq\lambda_2\ 矛盾。$$

从而可知，$\alpha+\beta$ 不是A的特征向量。

例6 设

$$A=\begin{pmatrix}1&-2&2\\-2&-2&4\\2&4&-2\end{pmatrix}.$$

计算 A^k ($k>0$)。

这个例子，直接计算 A^k 显然不行，那么若A相似于一个对角形就好办了。于是看A是否相似于对角形？

解 由

$$|\lambda I-A|=\begin{vmatrix}\lambda-1&2&-2\\2&\lambda+2&-4\\-2&-4&\lambda+2\end{vmatrix}=(\lambda-2)^2(\lambda+7)$$

知A的特征根为 2，2，-7。

而 $S_A(2)=\{k_1(2,0,1)+k_2(0,1,1)|\ k_1,k_2\in D\}$,

$S_A(-7)=\{k(1,2,-2)|\ k\in D\}$。

于是A的完全特征向量系为

$$(2,0,1),\ (0,1,1),\ (1,2,-2),$$

故

$$P=\begin{pmatrix}2&0&1\\0&1&2\\1&1&-2\end{pmatrix}.$$

则

$$P^{-1}AP=\begin{pmatrix}2&&\\&2&\\&&-7\end{pmatrix}.$$

因此

$$A^k=\left(P\begin{pmatrix}2&&\\&2&\\&&-7\end{pmatrix}P^{-1}\right)^k$$

$$=P\begin{pmatrix}2&&\\&2&\\&&-7\end{pmatrix}^kP^{-1}$$

$$=\begin{pmatrix}2&0&1\\0&1&2\\1&1&-2\end{pmatrix}\begin{pmatrix}2^k&&\\&2^k&\\&&(-7)^k\end{pmatrix}\begin{pmatrix}\frac{4}{9}&-\frac{1}{9}&\frac{1}{9}\\-\frac{2}{9}&\frac{5}{9}&\frac{4}{9}\\\frac{1}{9}&\frac{2}{9}&-\frac{2}{9}\end{pmatrix}$$

$$=\frac{1}{9}\begin{pmatrix}2^{k+3}+(-7)^k&-2^{k+1}+2(-7)^k&2^{k+1}-2(-7)^k\\-2^{k+1}+2(-7)^k&5\times2^k+4(-7)^k&2^{k+2}-4(-7)^k\\2^{k+1}-2(-7)^k&2^{k+2}-4(-7)^k&5\times2^k+4(-7)^k\end{pmatrix}.$$

例 7　证明：若 $A\sim B$，则A与B的最小多项式相同。反之，若A与B的最小多项式相同，那么A与B相似吗？

证明　设 $\varphi(\lambda)$，$\psi(\lambda)$ 分别为 A,B 的最小多项式，由于它们首系数为 1，若证它们相等，只须证互相整除即可。令$B=P^{-1}AP$。因为

$$\varphi(B)=\varphi(P^{-1}AP)=P^{-1}\varphi(A)P=0,$$

所以 $\psi(\lambda)|\varphi(\lambda)$。又因为

$$\psi(A)=\psi(PBP^{-1})=P\psi(B)P^{-1}=0,$$

所以 $\varphi(\lambda)|\psi(\lambda)$。因此 $\varphi(\lambda)=\psi(\lambda)$。

反之，不一定。比如

$$A=\begin{pmatrix}2&&\\&3&\\&&3\end{pmatrix},\quad B=\begin{pmatrix}2&&\\&2&\\&&3\end{pmatrix}$$

的最小多项式均为 $(\lambda-2)(\lambda-3)$，但 A与B不相似，因为A的特

征多项式为 $(\lambda-2)(\lambda-3)^2$与B的特征多项式$(\lambda-2)^2(\lambda-3)$不同。

例 8　设 e_1，e_2，e_3 为三维空间的标准正交基底，而

$$\alpha_1=\frac{1}{3}(2e_1+2e_2-e_3),\quad \alpha_2=\frac{1}{3}(2e_1-e_2+2e_3),$$

$$\alpha_3=\frac{1}{3}(e_1-2e_2-2e_3).$$

那么 α_1，α_2，α_3 也是标准正交基底。

证明　在第四节自学指导的说明 1）指出：一个向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 为标准正交基底必要而且只要

$$\alpha_i\alpha_j'=\begin{cases}1, & i=j\\ 0, & i\neq j.\end{cases}$$

于是下面只须证明这个条件成立即可。

事实上，
$$\begin{aligned}\alpha_1\alpha_1'&=\frac{1}{3}(2e_1+2e_2-e_3)\cdot\frac{1}{3}(2e_1+2e_2-e_3)'\\&=\frac{1}{9}(2e_1+2e_2-e_3)(2e_1'+2e_2'-e_3')\\&=\frac{1}{9}(4e_1e_1'+4e_1e_2'-2e_1e_3'+4e_2e_1'+4e_2e_2'\\&\quad-2e_2e_3'-2e_3e_1'-2e_3e_2'+e_3e_3')\\&=\frac{1}{9}(4+0-0+0+0+4-0-0-0+1)\\&=1.\end{aligned}$$

同理可证 $\alpha_1\alpha_2'=0$，$\alpha_1\alpha_3'=0$，$\alpha_2\alpha_1'=0$，$\alpha_2\alpha_2'=1$，$\alpha_2\alpha_3'=0$，$\alpha_3\alpha_1'=0$，$\alpha_3\alpha_2'=0$，$\alpha_3\alpha_3'=1$。故 α_1，α_2，α_3 为标准正交基底。

例 9　设A为正交阵，若 $|A|=-1$，则 -1必为A的特征根。

证明　要证-1是A的特征根，据定义，只须证 $|(-1)I-A|=0$ 即可。

事实上，
$$\begin{aligned}|(-1)I-A|&=|(-1)AA'-A|=|A((-1)A'-I)|\\&=|A|\ |(-1)A'-I|=|A|\ |-I-A'|\\&=|A|\ |(-I-A)'|=|A|\ |(-I-A)|,\end{aligned}$$

即
$$|(-1)I-A|=|A|\ |-I-A|.$$

由 $|A|=-1$，所以有

$$|(-1)I-A|=-|(-1)I-A|。$$

故

$$2|(-1)I-A|=0，\text{从而}\ |(-1)I-A|=0.$$

因此，-1是A的特征根。

例10 设A为 n阶实对称阵，且 $A^2=I_n$。证明，存在正交阵T，使

$$T^{-1}AT=\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&-1&&\\&&&&\ddots&\\&&&&&-1\end{pmatrix}.$$

证明 因A为实对称阵，据本章§5定理 2 的推论 1，存在正交阵T_1，使

$$T_1^{-1}AT_1=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}. \tag{1}$$

其中 λ_i 为A的特征根，因而都是实数。

又

$$T_1^{-1}AT_1T_1^{-1}AT_1=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix}=\begin{pmatrix}\lambda_1^2&&\\&\ddots&\\&&\lambda_n^2\end{pmatrix},$$

而

$$T_1^{-1}AT_1T_1^{-1}AT_1=T_1^{-1}AAT_1=T_1^{-1}A^2T_1=I_n.$$

所以，有

$$\begin{pmatrix}\lambda_1^2&&\\&\ddots&\\&&\lambda_n^2\end{pmatrix}=\begin{pmatrix}1&&\\&\ddots&\\&&1\end{pmatrix}$$

于是 $\lambda_i^2=1$，$\lambda_i=1$ 或-1，$i=1, 2, \cdots, n$。

把（1）中的 λ_i 适当的调换一下次序，使等于1的 λ_i 都调换到左上方，等于 -1 的 λ_i 均调到右下方的位置．这个调换过程，实际上调换 T_1 的列，而 T_1 是正交阵，交换列后仍为正交阵．令交换列后的正交阵为 T，因此有

$$T^{-1}AT=\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&-1&&\\&&&&\ddots&\\&&&&&-1\end{pmatrix}.$$

例11　设 $f(x_1, \cdots, x_n)=X'AX$ 是一个实二次型，$\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 的 n 个特征根，且 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$．证明，对任一实 n 维向量

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

均有

$$\lambda_1X'X\leqslant X'AX\leqslant\lambda_nX'X.$$

证明　由题设知 A 为对称阵，故存在正交阵 T，使

$$T'AT=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{pmatrix},$$

其中每个 λ_i 均为 A 的特征根．

在二次型 $f(x_1, \cdots, x_n)=X'AX$ 中，令满秩变数变换为

$$X=TY，其中\quad Y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}.$$

则有

$$f(x_1, \cdots, x_n)=X'AX=Y'T'ATY$$

$$=(y_1\cdots y_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}\begin{pmatrix}y_1\\ \vdots \\ y_n\end{pmatrix}$$

$$=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2.$$

于是

$$\lambda_1 Y'Y=\lambda_1(y_1^2+\cdots+y_n^2)\leqslant\lambda_1 y_1^2+\cdots+\lambda_n y_n^2\leqslant \lambda_n(y_1^2+\cdots+y_n^2)=\lambda_n Y'Y.$$

即 $\quad \lambda_1 Y'Y\leqslant X'A'X\leqslant\lambda_n Y'Y.$

而 $\quad X'X=Y'T'TY=Y'Y.$

所以 $\quad \lambda_1 X'X\leqslant X'AX\leqslant\lambda_n X'X.$

例12　设 A,B 均为正交阵．证明：存在正交阵T，使$T^{-1}AT=B$ 的充分必要条件是A与B的特征多项式相同．

证明　必要性是明显的．下面只须证明充分性．因为 A、B 均为正交阵，那么 A,B 的标准形均由各自的特征根唯一决定，充其量只差对角线上各块的顺序．由题设知A与B的特征根相同，故A与B的标准形相同，于是存在正交阵T_1，T_2，使

$$T_1^{-1}AT_1=T_2^{-1}BT_2 \text{ 或 } T_2T_1^{-1}AT_1T_2^{-1}=B.$$

令 $T=T_1T_2^{-1}$，则 $T^{-1}AT=B$.

例13　设

$$A=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}.$$

求 A^{10}

解　求 $\lambda I_3-A=\begin{pmatrix}\lambda+1 & -1 & 0\\ 4 & \lambda-3 & 0\\ -1 & 0 & \lambda-2\end{pmatrix},$

相抵标准形为

$$\begin{pmatrix} 1 & & \\ & 1 & \\ & & (\lambda-1)^2(\lambda-2) \end{pmatrix}.$$

于是A的初等因子组为 $\lambda-2$，$(\lambda-1)^2$，故

$$A \sim J = \begin{pmatrix} 2 & & \\ & 1 & 1 \\ & 0 & 1 \end{pmatrix}.$$

为了计算 A^{10} 必须计算出适合

$$P^{-1}AP = J$$

的可逆阵P。下面就这个例子介绍求P的方法：

假设

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & 1 \\ & & \lambda_2 \end{pmatrix}, \qquad P = (X_1 \ X_2 \ X_3).$$

那么

$$AP = P\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & 1 \\ & & \lambda_2 \end{pmatrix},$$

即

$$A(X_1 \ X_2 \ X_3) = (X_1 \ X_2 \ X_3)\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & 1 \\ & & \lambda_2 \end{pmatrix}.$$

亦即

$$(AX_1 \ AX_2 \ AX_3) = (\lambda_1 X_1 \ \lambda_2 X_2 \ X_2 + \lambda_2 X_3).$$

于是

$$AX_1 = \lambda_1 X_1, \quad AX_2 = \lambda_2 X_2, \quad AX_3 = X_2 + \lambda_2 X_3.$$

因此 $(\lambda_1 I - A)X_1 = 0$，$(\lambda_2 I - A)X_2 = 0$，$(\lambda_2 I - A)X_3 = -X_2$。

故 X_1，X_2 分别是A的属于 λ_1，λ_2 的特征向量，X_3 是上述最后一个方程组的解向量。把 X_1，X_2，X_3 求出后，也就求得了我们所

需要的P。

又假如

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & 1 & \\ & \lambda_1 & 1\\ & & \lambda_1\end{pmatrix},\qquad P=(X_1\ X_2\ X_3).$$

那么

$$A(X_1\ X_2\ X_3)=(X_1\ X_2\ X_3)\begin{pmatrix}\lambda_1 & 1 & \\ & \lambda_1 & 1\\ & & \lambda_1\end{pmatrix}.$$

于是

$$(AX_1\ AX_2\ AX_3)=(\lambda_1X_1\ X_1+\lambda_1X_2\ X_2+\lambda_1X_3).$$

因此

$$AX_1=\lambda_1X_1,\ AX_2=X_1+\lambda_1X_2,\ AX_3=X_2+\lambda_1X_3,$$

即

$$(\lambda_1I-A)X_1=0,\ (\lambda_1I-A)X_2=-X_1,\ (\lambda_1I-A)X_3=-X_2.$$

故X_1是A的属于 λ_1 的特征向量，解后两个线性方程组就得到X_2、X_3，从而就求得了所需要的P。

因为已经知道 P 是存在的，所以上面的后两个方程组一定有解，就是说 X_1，X_2，X_3 一定能求出来。当然 X_1，X_2，X_3 都不是唯一的，这也说明P不是唯一的。

在一般情况下，P也可以这样求得。

现在回过来求例13的P。已知A的特征根为 2,1,1. 而A的属于2的特征向量就是方程组

$$(2I-A)X=0\ \text{或}\ \begin{pmatrix}-2 & -6 & 0\\ 3 & 7 & 0\\ 3 & 6 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

的非零解，解之得一个解为

$$X_1=\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

又A的属于1的特征向量是方程组

$$(1\cdot I-A)X=0\text{ 或 }\begin{pmatrix}-3&-6&0\\3&6&0\\3&6&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

的非零解，解之得一解为

$$X_2=\begin{pmatrix}1\\2\\-1\end{pmatrix}.$$

再解方程组

$$(1\cdot I-A)X=X_2\text{ 或 }\begin{pmatrix}-3&-6&0\\3&6&0\\3&6&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1\\2\\-1\end{pmatrix},$$

得解为

$$X_3=\begin{pmatrix}0\\1\\-1\end{pmatrix}.$$

因此，所求的

$$P=\begin{pmatrix}0&1&0\\0&2&1\\1&-1&-1\end{pmatrix},\quad P^{-1}=\begin{pmatrix}-1&1&1\\1&0&0\\-2&1&0\end{pmatrix}.$$

故

$$A=PJP^{-1}=\begin{pmatrix}0&1&0\\0&2&1\\1&-1&-1\end{pmatrix}\begin{pmatrix}2&&\\&1&1\\&&1\end{pmatrix}\begin{pmatrix}-1&1&1\\1&0&0\\-2&1&0\end{pmatrix}$$

而

$$A^{10}=PJ^{10}P^{-1}=\begin{pmatrix}0&1&0\\0&2&1\\1&-1&-1\end{pmatrix}\begin{pmatrix}2^{10}&&\\&1&10\\&&1\end{pmatrix}\begin{pmatrix}-1&1&1\\1&0&0\\-2&1&0\end{pmatrix}$$

$$=\begin{pmatrix}0&1&10\\0&2&21\\2^{10}&-1&-11\end{pmatrix}\begin{pmatrix}-1&1&1\\1&0&0\\-2&1&0\end{pmatrix}$$

$$=\begin{pmatrix}-19&10&0\\-40&21&0\\-1003&-1035&1024\end{pmatrix}.$$

例14　证明：对任一λ-阵 $A(\lambda)$，都有

$$[D_k(\lambda)]^2 \mid D_{k-1}(\lambda)\cdot D_{k+1}(\lambda).$$

证明　因为

$$D_{k-1}(\lambda)=d_1(\lambda)\cdots d_{k-1}(\lambda)$$
$$D_k(\lambda)=d_1(\lambda)\cdots d_{k-1}(\lambda)d_k(\lambda)$$
$$D_{k+1}(\lambda)=d_1(\lambda)\cdots d_{k-1}(\lambda)d_k(\lambda)d_{k+1}(\lambda),$$

且令　$d_{k+1}(\lambda)=f(\lambda)d_k(\lambda)$.

于是

$$\begin{aligned}D_{k-1}(\lambda)D_{k+1}(\lambda)&=d_1^2(\lambda)\cdots d_{k-1}^2(\lambda)d_k(\lambda)d_{k+1}(\lambda)\\&=d_1^2(\lambda)\cdots d_{k-1}^2(\lambda)d_k^2(\lambda)f(\lambda)\\&=D_k^2(\lambda)f(\lambda).\end{aligned}$$

故　$D_k^2(\lambda)\mid D_{k-1}(\lambda)D_{k+1}(\lambda)$.

例15　设 n阶方阵A，适合 $A^2+A=2I_n$，试问A能否与对角形阵相似？

解　由$A^2+A=2I_n$，得$A^2+A-2I_n=0$，故A为多项式 $f(\lambda)=\lambda^2+\lambda-2$ 的根．而 $f(\lambda)=\lambda^2+\lambda-2=(\lambda-1)(\lambda+2)$，即 $f(\lambda)$ 没有重根．据本章§3命题 4 知，A的最小多项式 $\varphi(\lambda)$ 必整除 $f(\lambda)$，因此 $\varphi(\lambda)$ 无重根，于是A与对角形阵相似．

例16　求出三阶幂等阵A的一切可能的若当标准形．

解　由于 $A^2=A$，即 $A^2-A=0$，故A为多项式 $f(\lambda)=\lambda^2-\lambda$

的根．而A的最小多项式$\varphi(\lambda)$，$\varphi(\lambda)|f(\lambda)$，于是$\varphi(\lambda)$可能为$\lambda$，$\lambda-1$，$\lambda^2-\lambda$，也就是$A$的最后一个不变因子$d_3(\lambda)$可能为$\lambda,\lambda-1$，$\lambda^2-\lambda$，．

1）若$d_3(\lambda)=\lambda$，则必有$d_1(\lambda)=\lambda$，$d_2(\lambda)=\lambda$．这时A的初等因子组为λ，λ，λ．故A的若当标准形为

$$J=\begin{pmatrix}0&&\\&0&\\&&0\end{pmatrix};$$

2）若$d_3(\lambda)=\lambda-1$，则必有$d_2(\lambda)=\lambda-1$，$d_1(\lambda)=\lambda-1$．于是A的初等因子组为$\lambda-1$，$\lambda-1$，$\lambda-1$．故A的若当标准形为

$$J=\begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix};$$

3）若$d_3(\lambda)=\lambda^2-\lambda$，则$d_2(\lambda)=\lambda$，$d_1(\lambda)=1$或$d_2(\lambda)=\lambda-1$，$d_1(\lambda)=1$．

当$d_2(\lambda)=\lambda$时，A的初等因子组为λ，λ，$\lambda-1$，故A的若当标准形为

$$J=\begin{pmatrix}0&&\\&0&\\&&1\end{pmatrix};$$

当$d_2(\lambda)=\lambda-1$时，A的初等因子组为λ，$\lambda-1$，$\lambda-1$，故A的若当标准形为

$$J=\begin{pmatrix}0&&\\&1&\\&&1\end{pmatrix};$$

综上所述，三阶幂等阵A的一切可能的若当标准形有且只有下列四种：

$$\begin{pmatrix}0&&\\&0&\\&&0\end{pmatrix},\ \begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix},\ \begin{pmatrix}0&&\\&0&\\&&1\end{pmatrix},\ \begin{pmatrix}0&&\\&1&\\&&1\end{pmatrix}.$$

例17　设 n 阶方阵 A 的最小多项式为 $\varphi(\lambda)$，特征多项式为 $f(\lambda)$。证明 $f(\lambda)\mid(\varphi(\lambda))^n$。

证明　由于方阵A的最后一个不变因子是A的最小多项式，即 $d_n(\lambda)=\varphi(\lambda)$。而其他不变因子 $d_i(\lambda)$ $(i=1,\ 2,\ \cdots,\ n-1)$ 都能整除 $d_n(\lambda)$，也就是

$$d_i(\lambda)\mid\varphi(\lambda),\quad i=1,\ 2,\ \cdots,\ n-1.$$

从而有

$$d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda)\mid(\varphi(\lambda))^n.$$

而　$$f(\lambda)=|\lambda I-A|=d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda).$$

故　$$f(\lambda)\mid(\varphi(\lambda))^n.$$

第十一章　线性空间

〔内容提要〕

线性空间理论是以下三章的基础，也是高等代数的重要组成部分。

我们以n维向量空间——n数组空间$F^{(n)}$为基本模型，集中了分散着的某些高等代数的讨论对象，如$F[x]$，$M_n(F)$等，从中归纳出它们共同具有的基本性质，这样就导出数域F上线性空间的概念。

全章共七节，前四节讨论一般线性空间的基本性质，五、六两节介绍有限维空间的特殊的一些结果；最后一节给出线性空间的同构概念及简单性质。

本章的主要概念是：

线性空间、子空间、有限生成子空间；

子空间的交、和、直和；

线性表出与等价、线性相关与线性无关．极大线性无关组与秩；

有限维线性空间、基底，维数；

一个向量的坐标、一组向量的坐标矩阵；

同构映射、同构的线性空间。

本章的主要结果是：

1）非空子集S构成子空间的充分必要条件；

2）和空间是直和的充分必要条件；

3) 关于线性表出的基本性质，即§4命题3；

4) 极大线性无关组的存在性，即§4命题4；

5) n维线性空间中n个向量构成基底的条件，即§5命题1；

6) 有限维子空间的和与交的维数关系，即§5命题5；

7) 在有限维空间里，关于线性关系的向量、数组、矩阵的等价表现形式，即§6定理及推论；

8) 基底变换公式与坐标变换公式；

9) 同构的空间、同构映射的主要意义，即§7命题1；

10) 关于有限维空间的同构定理，即§7命题2；

11) 线性空间的同构关系是等价关系。

本章的主要方法，我们只强调一点，即讨论有限维空间的基本方法——坐标法。这就是在n维空间中取定一个基底之后，这个n维空间就完全的具体的转化为n数组空间$F^{(n)}$，而且把$F^{(n)}$中的数组写成$n\times 1$或$1\times n$矩阵的形式都可以。因此，n 维空间中一切与线性运算有关的问题都归结为n数组的情形去予以讨论、解决。

学习本章除了掌握好以上三个基本方面的内容之外，熟悉一批具体的例子也是非常必要的。正文各节中提供的例子已比较充分，首先应理解这批例子的内容、方法和意义。

〔内容分析〕

§1 线性空间的定义与简单性质

应当说，本节内容我们是比较熟悉的，因为在第六章已经相当充分地讨论过n维向量空间的基本性质。正因为这样，我们需要明确指出的是，线性空间概念的一般性。

首先，从名称上就有所不同。n维向量空间是专指n数组空间来说的，而线性空间的元素当然不一定是n数组。因此n维向量空间都是

线性空间，但一个线性空间不一定是n维向量空间。这说明，n维向量空间是线性空间的特例，而线性空间则是n维向量空间的一般化。线性空间概念所包含的范围要比n维向量空间广泛得多。不过，有的书上一律都用向量空间这一名称，这当然是可以的。

其次，应当注意到线性空间这一概念的抽象性。这主要体现在两个运算上。我们已经了解，一个集合上的运算叫什么名称并不是本质的东西。关键在于运算规则的特性。因此一个运算叫加法，未必就是我们通常所说的加法，退一步说，更不必是数的加法，因为施行运算的对象大可不必是数。同理，所说的倍数乘法当然也不一定是通常的倍数。这样，做为线性空间的两个运算：加法与倍数乘法，自然就不一定是通常意义下的加法与倍数乘法。之所以还这样叫法，就是因为还是用通常的加法与倍数乘法所具有的性质来要求这两个运算，即所说的八条运算性质。

现在我们用一个具体例子来说明一下这个问题。

D是实数域，D^+是正实数集。我们考虑以下两种运算

o_1：$(a,b)\longmapsto ab$，即$a o_1 b = ab$，$\forall a,b\in D^+$；

o_2：$(k,a)\longmapsto a^k$，即$k o_2 a = a^k$，$\forall k\in D, a\in D^+$。

显然，o_1这个运算就是通常数的乘法，o_2就是通常的指数运算。现在我们把o_1叫做“加法”，并且为了与这个叫法相适应，把o_1改记为$\oplus$；把o_2叫做倍数乘法，符号改记为$\odot$。于是我们可以验证，关于这两个运算：$\oplus$与$\odot$，D^+组成D上的线性空间。为此，必须验证八条算律都成立：

1） $a\oplus b = b\oplus a$

2） $(a\oplus b)\oplus c = a\oplus(b\oplus c)$

3） $a\oplus 1 = a = 1\oplus a$，$\forall a\in D^+$

4） $\forall a\in D^+$，有$a'\in D^+$使$a\oplus a' = 1 = a'\oplus a$

5） $k\odot(a\oplus b) = (k\odot a)\oplus(k\odot b)$

6） $(k+l)\odot a = (k\odot a)\oplus(l\odot a)$

7） $(kl)\odot a = k\odot(l\odot a)$

8) $1\odot a=a$.　　$k,l\in D$,　$a,b,c\in D^+$.

验证方法都一样，主要是记清$\oplus$与$\odot$的具体方法，比如验证5)成立：

$k\odot(a\oplus b)=k\odot(ab)=(ab)^k$,　　$(k\odot a)\oplus(k\odot b)=a^k\oplus b^k=a^kb^k=(ab)^k$

所以5）成立。这样，正实数集D^+就是实数域D上的线性空间。每一个正实数都是D^+中的向量，特别地，正实数1是D^+中的零向量，向量a的负向量是$\frac{1}{a}$。

§2 子空间

子空间概念在线性空间理论中有着重要的作用。

在子空间的定义里，明确要求子空间的运算与全空间的运算必须是一致的。这应当说是自然的，否则就失去子空间的意义了。

生成子空间的方法其重要意义在于它的普遍性，即任一子空间都可以用生成子空间的方法得到。

有限生成的空间是一类特殊而重要的线性空间，在高等代数里主要是讨论有限生成的线性空间。

§3 子空间的交与和　直和

子空间的交与和是比较具体的两个概念。值得注意的是，不要把两个子空间的和与两个子空间的并混同起来。二者有很大的差别。和是一个子空间，并则未必。并总是和的子集。

直和是个重要的概念，在以下的讨论中经常要用到，就本节来说，主要应当看到直和在分解化简线性空间方面的意义（命题1与3）。

§4 线性相关性

线性相关性概念是线性空间理论的基础．由于这些概念与 n 数组的情形完全一致，因此这里没有更多的话可说．

命题 3 是个很重要的结果，它体现了线性表出关系的一个基本特性．另外命题 3 的证明方法也是值得注意的，它说明，对一般线性空间的线性相关性问题，线性方程组的理论仍然起着重要的作用．

命题 4 及其推论的重要意义在于它对解决有限生成的线性空间的结构奠定了基础，这在以下两节中将看得很清楚．

命题 4 的证明可以用数学归纳法处理得清楚些．

事实上 $s=1$ 时，命题显然是成立的．

归纳假设对 $s-1$个向量的情形命题成立．去证对于s个向量的情形命题也成立．考虑

$$\alpha_1, \alpha_2, \cdots, \alpha_{s-1}, \alpha_s$$

的前 $s-1$个向量$\alpha_1, \alpha_2, \cdots, \alpha_{s-1}$．如果$\alpha_1, \alpha_2, \cdots, \alpha_{s-1}$ 全都是零向量，那么必有$\alpha_s \neq 0$．这时α_s就是$\alpha_1, \alpha_2 \cdots, \alpha_s$的极大线性无关组．如果$\alpha_1, \alpha_2, \cdots, \alpha_{s-1}$不全是零向量，由归纳假设，$\alpha_1, \alpha_2, \cdots, \alpha_{s-1}$ 有极大线性无关组．故可令

$$\alpha_{i_1}, \alpha_{i_2} \cdots, \alpha_{i_r}$$

是$\alpha_1, \alpha_2, \cdots, \alpha_{s-1}$的一个极大线性无关组．下面考察

$$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}, \alpha_s$$

如果它线性相关，那么$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}$就是$\alpha_1, \alpha_2, \cdots, \alpha_s$ 的极大线性无关组．如果$\alpha_{i_1}, \alpha_{i_2}, \cdots, \alpha_{i_r}, \alpha_s$是线性无关的，那么它就是$\alpha_1, \alpha_2, \cdots, \alpha_s$的一个极大线性无关组．

§5 有限维空间

基底和维数是有限维空间的基本特征．有限维空间理论的任何

一个结果都是通过基底和维数体现的。因此了解有限维空间的基底性质无疑是必要的。对此，命题 1 是有用的。

命题 3 说明有限生成的与有限维的这两个概念是等价的。

命题 5 是个重要结果，在许多关于子空间的问题中有着重要作用。

最后我们指出，只含一个向量的零空间 {0}，按定义，是有限维空间，维数等于零。但这个有限维空间没有基底。当然，有维数而没有基底的线性空间就此一个。

§6 坐　标

坐标进一步体现了有限维空间的本质．正因为有了坐标，我们才对这样或那样的有限维空间有了统一的，具体的，本质的认识。正是利用坐标，我们使得一般的向量转化为n数组，并且发现任一n维空间与n数组空间的一致性．还是利用坐标，使得矩阵理论成了解决有限维线性空间的线性相关性问题的有效工具。

十分明白，坐标是相对于一个基底来说的．因此，自然需要解决在不同基底上的坐标之间的关系。命题 1 与命题 2 回答了这个问题。

§7 线性空间的同构

同构是代数学中重要的概念。通俗地说，同构就是结构相同的意思。可是，对于一个线性空间来说，什么叫做它的结构呢？通过以上几节的讨论，我们已经比较充分的觉察到，对于线性空间$V(F)$来说，决定性的东西不是V中的元素如何，而是这些元素所施行的运算：加法与倍数。因此，由这两个运算所决定的性质才是做为线性空间的$V(F)$的本质属性。所以我们可以这样说，由$V(F)$的加法与倍数乘法这两个运算所决定的性质就体现了$V(F)$的结

构．本着这样的意思，线性空间的同构和同构映射的意义就比较易于理解了。

上述观念是掌握本节理论并能用来解决一些具体问题的基础。

〔例题选解〕

例 1　替换定理　假设

1）$\alpha_1,\alpha_2,\cdots,\alpha_r$线性无关，

2）$\alpha_1,\alpha_2,\cdots,\alpha_r$能用$\beta_1,\beta_2,\cdots,\beta_s$线性表出。那么可用$\alpha_1,\alpha_2,\cdots,\alpha_r$替换$\beta_1,\beta_2,\cdots,\beta_s$中的 r 个向量使

$$\alpha_1,\alpha_2\cdots,\alpha_r,\ \beta_{i_1},\ \beta_{i_2},\cdots,\beta_{i_{t-r}}$$

与$\beta_1,\beta_2,\cdots,\beta_s$等价。

证明　令$\beta_1,\beta_2,\cdots,\beta_s$的秩为　t，不妨设

$$\beta_1,\beta_2\cdots,\beta_t$$

是一个极大线性无关组，故$\beta_1,\beta_2,\cdots,\beta_t$ 与 $\beta_1,\beta_2,\cdots,\beta_s$等价，因而$\alpha_1,\alpha_2,\cdots,\alpha_r$能用$\beta_1,\beta_2\cdots,\beta_t$线性表出。下面做为一组向量来考察

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_1,\beta_2,\cdots,\beta_t \qquad (*)$$

$\alpha_1,\alpha_2,\cdots,\alpha_r$是（*）中一组线性无关向量，所以可以扩充成（*）的一个极大线性无关组如下：

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_{i_1},\cdots,\beta_{i_{t-r}}$$

于是$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_{i_1},\cdots,\beta_{i_{t-r}},\beta_{t+1},\cdots,\beta_s$ 就是用 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 替换$\beta_1,\cdots,\beta_t,\beta_{t+1},\cdots,\beta_s$中的$r$个向量之后得到的与$\beta_1,\beta_2,\cdots,\beta_s$ 等价的一组向量。

例 2　$W_1\cup W_2=W_1+W_2$当且仅当$W_1\subseteq W_2$或$W_2\subseteq W_1$。

解　若 $W_1\cup W_2=W_1+W_2$，$W_2\nsubseteq W_1$，　于是对任意的 $\alpha_1+\alpha_2\in W_1+W_2$，$\alpha_1\in W_1$，$\alpha_2\in W_2$，或者$\alpha_1+\alpha_2\in W_1$或$\alpha_1+\alpha_2\in W_2$。但是$\alpha_1+\alpha_2\notin W_1$否则$\alpha_2\in W_1$从而$W_2\subseteq W_1$这是不可能的。所以，$\alpha_1+\alpha_2\in W_2$，因而$\alpha_1\in W_2$，即$W_1\subseteq W_2$。

反之，$W_1\subseteq W_2$ 时，$W_1\cup W_2=W_2=W_1+W_2$；$W_2\subseteq W_1$ 时，

$W_1 \cup W_2 = W_1 = W_1 + W_2$。

例 3　和$W_1 + W_2 + \cdots + W_s$是直和的充分必要条件是 $(W_1 + \cdots + W_{i-1}) \cap W_i = \{0\}$，$i = 2, 3, \cdots, s$。

证明　必要性是自明的。为了证明充分性，我们指出零向量的表示式是唯一的。设

$$\alpha_1 + \alpha_2 + \cdots + \alpha_s = 0,\ \alpha_i \in W_i$$

如果α_i不全为零，必有$\alpha_r \neq 0$，$i > r$ 时　$\alpha_i = 0$。于是

$$\alpha_1 + \alpha_2 + \cdots + \alpha_r = 0,\ \alpha_r \neq 0.$$

从而

$$\alpha_r = -(\alpha_1 + \cdots + \alpha_{r-1}).$$

这表明 $\alpha_r \in (W_1 + \cdots + W_{r-1}) \cap W_r$。这是个矛盾。所以零向量的表示法是唯一的，故充分性得证。

例 4　在$M_n(F)$中求子空间

$$W = \{(a_{ij}) \mid a_{11} + a_{22} + \cdots + a_{nn} = 0\}$$

的维数。

解　令$W_1 = \{(a_{ij}) \in W \mid a_{ii} = 0\}$，$W_2 = \{(a_{ij}) \in W \mid i \neq j$ 时$a_{ij} = 0\}$。于是 $W = W_1 \dot{+} W_2$。　$\dim W_1 = n^2 - n$，　$\dim W_2 = n - 1$。所以 $\dim W = (n^2 - n) + (n - 1) = n^2 - 1$。

例 5　线性空间V只有有限个不同的子空间充分必要条件是 $\dim V < 2$。

证明　充分性是明显的。下面证明必要性。用反证法。假设 $\dim V \geqslant 2$。从而有$\alpha, \beta \in V$，但α, β 是线性无关的。考察以下类型的一维子空间：

$$L(\alpha + r\beta)\ 与\ L(\alpha + s\beta).$$

我们证明，只要$r \neq s$，必有$L(\alpha + r\beta) \neq L(\alpha + s\beta)$，否则可得 $\alpha + r\beta = k(\alpha + s\beta)$，并且$k \neq 1$。于是则有

$$(k-1)\alpha = (r - ks)\beta$$

因为$k \neq 1$，上式说明α与β是线性相关的。这个矛盾肯定了$L(\alpha + n\beta)$，$n \in N$，是两两不同的子空间，故V有无限多个不同的子空间。

第十二章　线性变换

〔内容提要〕

如正文所述，线性变换的概念就是线性函数的发展和一般化．通过大量的实例可以看出线性变换概念的普遍性，不仅是代数学的重要讨论对象，在其它的数学课程里也有广泛的应用．

全章七节分成两段，前四节讨论一般线性空间上的线性变换，后三节专门讨论有限维空间上的线性变换．

本章的主要概念是

线性变换、线性变换的运算，可逆变换；

不变子空间、导出线性变换，值域与核；

线性变换的特征向量、特征值；

线性变换的特征子空间；

线性变换的表示矩阵、标准表示矩阵；

线性变换的秩与亏数．

本章的主要结果是

1．在一个线性变换 σ 之下，子空间的象是子空间，子空间的原象是子空间．

2．不变子空间的交与和还是不变子空间．

3．特征向量的性质，即§4命题２、命题３．

4．n维空间的线性变换与n阶方阵的关系，即§5定理及命题１．

5．秩与亏数的关系，即§5命题３．

6．线性变换的特征向量、特征值与其表示矩阵的特征根、特

征向量的一致性，即§5命题4.

7. 线性变换在各个基底上表示矩阵之间的关系，即 §6 定理1、定理2.

8. 线性变换的有理标准形的表示矩阵和若当标准形的表示矩阵，即§7定理1、定理2.

本章的主要方法是

1. 求线性变换的特征根、特征向量和特征子空间. 这主要是对有限维空间来说的，此时问题完全归结为求表示矩阵的特征根、特征向量和特征子空间.

2. 求标准表示矩阵. 这相当于求相应的相似标准形.

与上一章一样，除了掌握好以上三个基本方面的内容之外，熟悉一些具体的例子也是很必要的. 正文中讲到的例题已为数不少，应尽量先理解好这些例题中的内容方法和意义.

〔内容分析〕

关于线性变换的定义、性质和运算都是容易理解的. 因为这些事实的背景是很具体的. 也是我们比较熟悉的. 有一点值得提出的是线性变换的乘法. 我们知道，通常两个函数相乘和相加一样，是通过函数值的相乘、相加来实现的，线性变换的加法就是这样，由象相加来规定的. 但线性变换的乘法则不然. 它不是用象相乘来规定的. 这固然是因为，一般来说线性空间中的向量都没规定乘法，从而无从说用象（向量）相乘来规定变换的乘法. 其实，既使向量可以相乘，我们也不用它来规定线性变换的乘法. 道理在于，用函数值相乘定义函数乘法时，线性的性质已不能保持，即线性函数相乘得到的函数不再是线性的. 这当然是我们所不希望的. 从而，我们按复合函数的方式来定义线性变换的乘法. 当然这对我们也是很习惯的.

在前两节中，线性变换运算所满足的许多性质也是值得提及

的，特别地，线性空间$V(F)$上一切线性变换组成的集合$L(V)$，关于线性变换的加法和数域F的数与线性变换的倍数乘法来说，完全具备了线性空间所要求的一切条件，即 $L(V)$构成一个F上的线性空间。

三、四两节讨论了线性变换与子空间的关系，中心意思在于揭示一个子空间在线性变换下的变动状态。这在正文中已有清楚的叙述。

值得明确强调的是不变子空间及导出线性变换在分解化简线性变换方面的作用和意义。正文有关这部分的文字读者应当多加注意加深理解。

最后提出的与一维不变子空间密切相关的特征向量、特征值概念是很重要的两个概念。进而给出特征子空间的定义及重要性质。所有这些在化简线性变换的理论中起着特别重要的作用。

有限维空间的线性变换理论是高等代数的极其重要的课题，也是相当精彩的部分。它的基本特征就是抽象对象具体化，使用的手段是我们早已熟悉的坐标法。

这一节的主要结果全都体现在正文中的定理和四个命题与推论上这些结果都是非常重要的，是讨论有限维空间线性变换问题主要的基础、常用的工具。因此应当深入理解、牢固掌握和灵活运用它们去解决一些其它问题。

线性变换的表示矩阵，与向量的坐标一样，是相对于一个基底来说的。因此，同一个线性变换在各个基底上的一切表示矩阵是一个矩阵的子集。这样自然提出的问题是这样的矩阵子集有何特性。第六节的定理 1 与定理 2 很好的回答了这个问题。结论是：同一个线性变换在各个基底上的一切表示矩阵组成的矩阵子集恰好是一个相似类。

在此基础上，进而就提出线性变换的标准表示矩阵的问题。这个问题就是相似矩阵的标准形问题。从而，把我们在第十章得到的 n 阶方阵的有理标准形定理和若当标准形定理转述过来就解决了问

题．这就是第七节的定理1和定理2．

〔例题选解〕

例1　设σ是线性空间V的线性变换．证明：

1）$\sigma(V)\subseteq\sigma^{-1}(0)$当且仅当$\sigma^2=0$，

2）$\sigma^{-1}(0)\subseteq(\sigma^2)^{-1}(0)\subseteq(\sigma^3)^{-1}(0)\subseteq\cdots$，

3）$\sigma(V)\supseteq\sigma^2(V)\supseteq\sigma^3(V)\supseteq\cdots$．

证明 1）若$\sigma(V)\subseteq\sigma^{-1}(0)$，任取$\alpha\in V$，则$\sigma(\alpha)\in\sigma^{-1}(0)$，从而$\sigma^2(\alpha)=\sigma(\sigma(\alpha))=0$，即$\sigma^2=0$．

反之若$\sigma^2=0$．$\beta\in\sigma(V)$，则有$\alpha\in V$使$\beta=\sigma(\alpha)$．于是$\sigma(\beta)=\sigma(\sigma(\alpha))=\sigma^2(\alpha)=0$．故$\beta\in\sigma^{-1}(0)$，即$\sigma(V)\subseteq\sigma^{-1}(0)$．

2）这是比较明显的．因为对任意正整数k，如果$\sigma^k(\alpha)=0$．即$\alpha\in(\sigma^k)^{-1}(0)$，都有$\sigma^{k+1}(\alpha)=\sigma(\sigma^k(\alpha))=0$，即$\alpha\in(\sigma^{k+1})^{-1}(0)$，所以对任一$k\in N$，$(\sigma^k)^{-1}(0)\subseteq(\sigma^{k+1})^{-1}(0)$．

3）任一$\beta\in\sigma^{k+1}(V)$，则有$\alpha\in V$使$\beta=\sigma^{k+1}(\alpha)$．于是$\beta=\sigma^{k+1}(\alpha)=\sigma^k(\sigma(\alpha))\in\sigma^k(V)$，所以$\sigma^{k+1}(V)\subseteq\sigma^k(V)$，$k\in N$．

例2　设σ是线性空间V的一个线性变换，并且$\sigma^2=\sigma$．证明：

1）$\sigma^{-1}(0)=\{\alpha-\sigma(\alpha)\mid\alpha\in V\}$，

2）$V=\sigma^{-1}(0)\oplus\sigma(V)$，

3）如果τ是V的一个线性变换，那么$\sigma^{-1}(0)$，$\sigma(V)$是τ的不变子空间的充分必要条件是$\sigma\tau=\tau\sigma$．

证明1）对任一$\alpha\in V$，$\sigma(\alpha-\sigma(\alpha))=\sigma(\alpha)-\sigma^2(\alpha)=0$，即$\alpha-\sigma(\alpha)\in\sigma^{-1}(0)$．反之若$\beta\in\sigma^{-1}(0)$，即$\sigma(\beta)=0$，于是$\beta=\beta-\sigma(\beta)$．这样$\sigma^{-1}(0)=\{\alpha-\sigma(\alpha)\mid\alpha\in V\}$．

2）对任一$\alpha\in V$，由$\alpha=\alpha-\sigma(\alpha)+\sigma(\alpha)$可知

$$V=\sigma^{-1}(0)+\sigma(V)．$$

我们指出，这是一个直和．假设$\beta\in\sigma^{-1}(0)\cap\sigma(V)$．于是$\sigma(\beta)=0$，且有$\alpha\in V$，使$\beta=\sigma(\alpha)$，从而

$$\beta = \sigma(\alpha) = \sigma^2(\alpha) = \sigma(\sigma(\alpha)) = \sigma(\beta) = 0.$$

所以$\sigma^{-1}(0) \cap \sigma(V) = \{0\}$，即$V = \sigma^{-1}(0) \oplus \sigma(V)$。

3）若$\sigma\tau = \tau\sigma$，对任一$\alpha \in \sigma^{-1}(0)$，则$\sigma(\alpha) = 0$，于是 $\sigma(\tau(\alpha)) = \sigma\tau(\alpha) = \tau\sigma(\alpha) = \tau(\sigma(\alpha)) = \tau(0) = 0$。即$\tau(\alpha) \in \sigma^{-1}(0)$。所以 $\sigma^{-1}(0)$是τ的不变子空间。

如果$\alpha \in \sigma(V)$，则$\alpha = \sigma(\beta)$．于是 $\tau(\alpha) = \tau(\sigma(\beta)) = \tau\sigma(\beta) = \sigma\tau(\beta) = \sigma(\tau(\beta)) \in \sigma(V)$，所以$\sigma(V)$是$\tau$的不变子空间。

反之，对任一$\alpha \in V$，则有

$$\begin{aligned}\sigma\tau(\alpha) &= \sigma\tau(\alpha - \sigma(\alpha) + \sigma(\alpha)) = \sigma(\tau(\alpha - \sigma(\alpha)) + \tau\sigma(\alpha)) \\ &= \sigma(\tau(\alpha - \sigma(\alpha))) + \sigma(\tau\sigma(\alpha))\end{aligned}$$

由于$\sigma^{-1}(0)$，$\sigma(V)$都是τ的不变子空间，则有

$$\tau(\alpha - \sigma(\alpha)) \in \sigma^{-1}(0),\quad \tau\sigma(\alpha) = \sigma(\beta),\quad \beta \in V.$$

从而即有

$$\sigma(\tau(\alpha - \sigma(\alpha))) = 0,\ \sigma(\tau\sigma(\alpha)) = \sigma(\sigma(\beta)) = \sigma^2(\beta) = \sigma(\beta) = \tau\sigma(\alpha)$$

由此使得$\sigma\tau(\alpha) = 0 + \tau\sigma(\alpha) = \tau\sigma(\alpha)$，所以$\sigma\tau = \tau\sigma$。

例 3　在$M_2(F)$中，取定$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$。令

$$\sigma:\ X \longmapsto AX,\ \forall X \in M_2(F),$$

证明：σ是$M_2(F)$的线性变换，并求$\sigma^{-1}(0)$、$\sigma(M_2(F))$。

证明　对任意的$X, Y \in M_2(F)$，$k, l \in F$，则有

$$A(kX + lY) = k(AX) + l(AY),$$

由此得出$\sigma(kX + lY) = k\sigma(X) + l\sigma(Y)$。即 σ 是 $M_2(F)$ 的线性变换。

显然，$\sigma^{-1}(0) = \{X \mid AX = 0\}$ 设

$$X = \begin{pmatrix} x_1 & x_3 \\ x_2 & x_4 \end{pmatrix}$$

于是

$$AX=0 \text{ 当且仅当} \begin{pmatrix} A & \\ & A \end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{pmatrix}=\begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}.$$

下面我们解齐次线性方程组

$$\begin{pmatrix} 1 & 2 & & \\ 2 & 4 & & \\ & & 1 & 2 \\ & & 2 & 4 \end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4 \end{pmatrix}=\begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}.$$

经计算得基础解系为$\alpha_1=(-2,1,0,0)$，$\alpha_2=(0,0,-2,1)$从而$X_1=\begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}$，$X_2=\begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix}$是$\sigma^{-1}(0)$的一个基底，即得

$$\sigma^{-1}(0)=\left[\begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix}\right]$$

最后，在$\alpha_1=(-2,1,0,0,)$，$\alpha_2=(0,0,-2,1)$的基础上扩充成$F^{(4)}$的基底. 如添上$\alpha_3=(1,0,0,0)$，$\alpha_4=(0,0,0,1)$就得 $F^{(4)}$的一个基底，相应地，则有

$$X_1=\begin{pmatrix} -2 & 0 \\ 1 & 0 \end{pmatrix}, X_2=\begin{pmatrix} 0 & -2 \\ 0 & 1 \end{pmatrix}, X_3=\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$X_4=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

是$M_2(F)$的基底. 这样, AX_3，AX_4便是$\sigma(M_2(F))$的基底，即有

$$\sigma(M_2(F))=[AX_3, AX_4].$$

其中

$$AX_3=\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix},$$

$$AX_4=\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}0&2\\0&4\end{pmatrix},$$

所以

$$\sigma(M_2(F))=\left\{\begin{pmatrix}1&0\\2&0\end{pmatrix},\begin{pmatrix}0&2\\0&4\end{pmatrix}\right\}=\left\{\begin{pmatrix}1&0\\2&0\end{pmatrix},\begin{pmatrix}0&1\\0&2\end{pmatrix}\right\}$$

例 4 在$M_2(F)$中，σ如例 3，求σ在基底

$$\varepsilon_{11}=\begin{pmatrix}1&0\\0&0\end{pmatrix},\quad \varepsilon_{12}=\begin{pmatrix}0&1\\0&0\end{pmatrix},\quad \varepsilon_{21}=\begin{pmatrix}0&0\\1&0\end{pmatrix},$$

$$\varepsilon_{22}=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

上的表示矩阵．并求σ的秩与亏数．

解

$$\sigma(\varepsilon_{11})=\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}=\begin{pmatrix}1&0\\2&0\end{pmatrix}=\varepsilon_{11}+2\varepsilon_{21},$$

$$\sigma(\varepsilon_{12})=\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}0&1\\0&0\end{pmatrix}=\begin{pmatrix}0&1\\0&2\end{pmatrix}=\varepsilon_{12}+2\varepsilon_{22},$$

$$\sigma(\varepsilon_{21})=\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}0&0\\1&0\end{pmatrix}=\begin{pmatrix}2&0\\4&0\end{pmatrix}=2\varepsilon_{11}+4\varepsilon_{21},$$

$$\sigma(\varepsilon_{22})=\begin{pmatrix}1&2\\2&4\end{pmatrix}\begin{pmatrix}0&0\\0&1\end{pmatrix}=\begin{pmatrix}0&2\\0&4\end{pmatrix}=2\varepsilon_{12}+4\varepsilon_{22}.$$

从而σ的表示矩阵为

$$B=\begin{pmatrix}1&0&2&0\\0&1&0&2\\2&0&4&0\\0&2&0&4\end{pmatrix}.$$

显然，$\operatorname{rank}B=2$，所以σ的秩等于 2 ，从而σ的亏数也等于 2 。

第十三章　欧氏空间及其线性变换

〔内容提要〕

在特殊的线性空间，二、三维几何空间里，度量性质是它的重要特点，在许多问题上有突出的作用。把度量性质推广到一般的线性空间，从而提出了欧氏空间的概念。

本章分三节讲三个问题。

主要概念是

欧氏空间；

向量长、夹角、单位向量、正交；

标准正交向量组；

度量矩阵、标准正交基底，同构；

对称变换、正交变换。

主要结果是

1. 柯西不等式。
2. n 维线性空间构成欧氏空间的普遍方法。
3. 度量矩阵的基本性质。
4. 标准正交基底的存在性。
5. 有限维欧氏空间的同构定理。
6. 对称变换的基本性质及标准表示矩阵。
7. 正交变换的基本性质及标准表示矩阵。

主要方法是

1. 对任一 n 维线性空间 V 规定内积使之构成欧氏空间；用正

定矩阵规定内积使V构成欧氏空间。

2．求标准正交基底的两种方法：初等变换法和正交化法。

3．求对称变换、正交变换的标准表示矩阵。

〔内容分析〕

第一节的内容比较简单，其中提到的概念都很具体而且有相当的直观性。我们强调以下三点：

1）灵活掌握两个典型的例子，即n数组空间$D^{(n)}$和连续函数空间$C[a,b]$。

2）熟记柯西不等式。

3）对正交、向量长既要了解它的直观性一面，又要看到它的抽象性的一面。切不可把一般欧氏空间中的正交性与向量长和几何空间的直观现象混同起来。

第二节颇为重要。方法性强，概念性也强。本节有两个核心概念：度量矩阵与标准正交基底，围绕度量矩阵概念揭示出线性空间与欧氏空间之间的区别和联系，从而掌握了使线性空间构成欧氏空间的普遍方法。通过标准正交基底确定了欧氏空间的结构。有限维欧氏空间的唯一特征就是它的维数。

本节例2值得注意，它说明对同一个线性空间用不同的度量矩阵可以得到大不一样的欧氏空间。

第三节利用对称矩阵与正交矩阵定义了对称变换与正交变换。务必注意的是，这个定义依赖于标准正交基底。

第七章　集合与映射

练　习　一

1．1）若$A\cap B=A$，对任一$x\in A$，有$x\in A\cap B$．从而$x\in B$．所以$A\subseteq B$．反之，若$A\subseteq B$则$A\subseteq A\cap B$．又$A\cap B\subseteq A$，所以$A\cap B=A$．

2）同理可证．

2、取$C=A$，$D=\phi$即有　$B\cup A=A$，$B\cap\phi=\phi$．一般来说 C 与D不是唯一的．比如还可以取：$C=A-B$，$D=A-B$，那么也有$B\cup(A-B)=A$，$B\cap(A-B)=\phi$．而当$B\neq\phi$，$B\neq A$时，$A-B\neq A$，$A-B\neq\phi$．如 $A=\{1,2\}$，$B=\{1\}$，则有 $A-B=\{2\}$．如果要求$C=D$，即

$$B\cup C=A,\ B\cap C=\phi,$$

那么这样的C就是唯一的．即必须是$C=A-B$．因为由 $B\cup C=A$ 可知$C\subseteq A$，又由$B\cap C=\phi$，可得$C\subseteq A-B$．而$A-B\subseteq C$．所以$C=A-B$．

3．令$A=\{1\}$，$B=\{2\}$ 即可．此题说明，集合之间的包含关系$\subseteq$与数之间的大小关系$\leqslant$有类似的地方，但不完全一样，因为任意两个数a与b，对于小于关系$<$来说，以下三种情形有且只有一种成立：

$$a<b,\ a=b,\ b<a.$$

4．$A=\{a,b,c\}$共有八个子集如下：

$\{a\}$，$\{b\}$，$\{c\}$，$\{a,b\}$，$\{a,c\}$，$\{b,c\}$，$\{a,b,c\}$，ϕ．

所以$P(A)=\{\{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}, \phi\}$.

练 习 二

1　1) $\alpha=\rho(\cos\theta+i\sin\theta)\longmapsto(\rho+1)(\cos\theta+i\sin\theta)$,（当然$0\mapsto1$）;

2) $\alpha=\rho(\cos\theta+i\sin\theta)$　当$\rho>1$时规定

$$\rho(\cos\theta+i\sin\theta)\longmapsto(\rho-1)(\cos\theta+i\sin\theta),$$

当$\rho\leqslant1$时，规定

$$\rho(\cos\theta+i\sin\theta)\longmapsto\rho(\cos\theta+i\sin\theta).$$

3) $\alpha=a+bi\longmapsto a$;

4) $\alpha=a+bi\longmapsto\overline{\alpha}=a-bi$.

2. $\varphi: x\longmapsto10^x$.　φ的逆映射为

$$\psi:\ x\longmapsto\log_{10}x$$

3. φ: $n\longmapsto2n$, $n>0$;

$$n\longmapsto|2n|+1,\ n\leqslant0.$$

φ的逆映射为

$$\psi: 2n\longmapsto n,\ n>0,$$

$$|2n|+1\longmapsto n,\ n\leqslant0.$$

4. φ_1与φ_2都是$F[x]$的变换是明显的. 我们指出二者都是单满变换.

满的：$\forall g(x)\in F[x]$,

$$\varphi_1:\ g(x-h)\longmapsto g(x),$$

$$\varphi_2:\ g\left(\frac{1}{k}x\right)\longmapsto g(x).$$

单的：由$f_1(x+h)=f_2(x+h)$,可以得出$f_1(y)$与$f_2(y)$对y取任何值c时，都有$f_1(c)=f_2(c)$. 由此可得$f_1(y)$与$f_2(y)$是相同的两个多项式. 即$f_1(x)=f_2(x)$, 这就说明φ_1是单变换. 同理φ_2也是

单变换。

练　习　三

1．因为$A=\{-1,\ 1\}$，我们可以自然的得到A的两个代数运算：

$$\varphi_1:(a,b)\longmapsto ab,\quad \forall a,\ b\in A;$$

$$\varphi_2:(a,b)\longmapsto a^b,\quad \forall a,b\in A.$$

这里φ_1就是通常的乘法，所以代数运算φ_1满足结合律也满足交换律，而φ_2是以a为底的幂的运算，具体写出来就是：

$$(-1)\varphi_2(-1)=(-1)^{-1}=-1,\quad (-1)\varphi_2 1=(-1)^1=-1,$$

$$1\varphi_2(-1)=1^{-1}=1,\quad 1\varphi_2 1=1^1=1.$$

显然，φ_2不满足交换律，但是容易发现：总有$a\varphi_2 b=a$成立．从而

$$(x\varphi_2 y)\varphi_2 z=x\varphi_2 z=x.$$

$$x\varphi_2(y\varphi_2 z)=x\varphi_2 y=x.$$

所以φ_2满足结合律。

下面我们考察这两个运算的分配律问题．

$\because$　$x\varphi_1(y\varphi_2 z)=x\varphi_1 y,\quad (x\varphi_1 y)\varphi_2(x\varphi_1 z)=x\varphi_1 y$

$\therefore$　$x\varphi_1(y\varphi_2 z)=(x\varphi_1 y)\varphi_2(x\varphi_1 z).$

$\because$　$(y\varphi_2 z)\varphi_1 x=y\varphi_1 x,\quad (y\varphi_1 x)\varphi_2(z\varphi_1 x)=y\varphi_1 x$

$\therefore$　$(y\varphi_2 z)\varphi_1 x=(y\varphi_1 x)\varphi_2(z\varphi_1 x).$

这证明了运算φ_1对运算φ_2适合分配律。

$\because$　$(y\varphi_1 z)\varphi_2 x=y\varphi_1 z,\quad (y\varphi_2 x)\varphi_1(z\varphi_2 x)=y\varphi_1 z$

$\therefore$　$(y\varphi_1 z)\varphi_2 x=(y\varphi_2 x)\varphi_1(z\varphi_2 x).$

$\because$　$x\varphi_2(y\varphi_1 z)=x,\quad (x\varphi_2 y)\varphi_1(x\varphi_2 z)=x\varphi_1 x$

$\therefore$　当$x=-1$时　$x\varphi_2(y\varphi_1 z)\neq(x\varphi_2 y)\varphi_1(x\varphi_2 z).$

这就说明，运算φ_2对运算φ_1满足右分配律，但不满足左分配律。

2．对$A=\{0,1,2\}$来说，容易想到，仿照上题的做法就不行了。因此，应当另做考虑，对于0，1，2这三个数来说还有哪些

"两得一个"的办法。比如求大公约、小公倍的办法，对0，1，2这三个数组成的集合A来说就是"两得一个"的办法，即$\forall a,b\in A=\{0,1,2\}$，规定

$$a\varphi_1 b=(a,b),\quad a\varphi_2 b=[a,b].$$

其中(a,b)与$[a,b]$分别表示a与b的大公约与小公倍。具体写出φ_1,φ_2的运算结果就是：

$$0\varphi_1 0=(0,0)=0,\quad 0\varphi_1 1=(0,1)=1,\quad 0\varphi_1 2=(0,2)=2,$$

$$1\varphi_1 0=(1,0)=1,\quad 1\varphi_1 1=(1,1)=1,\quad 1\varphi_1 2=(1,2)=1,$$

$$2\varphi_1 0=(2,0)=2,\quad 2\varphi_1 1=(2,1)=1,\quad 2\varphi_1 2=(2,2)=2.$$

$$0\varphi_2 0=[0,0]=0,\quad 0\varphi_2 1=[0,1]=0,\quad 0\varphi_2 2=[0,2]=0,$$

$$1\varphi_2 0=[1,0]=0,\quad 1\varphi_2 1=[1,1]=1,\quad 1\varphi_2 2=[1,2]=2,$$

$$2\varphi_2 0=[2,0]=0,\quad 2\varphi_2 1=[2,1]=2,\quad 2\varphi_2 2=[2,2]=2.$$

这两个运算都满足交换律是自明的。它们都满足结合律也是早已知道的。即

$$((a,b),c)=(a,(b,c)),\quad [[a,b],c]=[a,[b,c]].$$

下面考虑这两个运算的分配律问题。我们分几种情形来证明，$\forall x,y,z\in A$都有

$$x\varphi_1(y\varphi_2 z)=(x\varphi_1 y)\varphi_2(x\varphi_1 z).$$

1）x,y,z中只少有一个是0．$x=0$时，等式两端都等于$y\varphi_2 z$；$y=0$（$z=0$）时，等式两端都等于x。

2）x,y,z 中只少有一个是1．

（i）$x=1$ 时

$$1\varphi_1(y\varphi_2 z)=1,\quad (1\varphi_1 y)\varphi_2(1\varphi_1 z)=1\varphi_2 1=1$$

所以等式成立。

(ii) $y=1$ 时

$$x\varphi_1(1\varphi_2 z)=x\varphi_1 z,\quad (x\varphi_1 1)\varphi_2(x\varphi_1 z)=x\varphi_1 z$$

所以等式成立。

(iii) $z=1$ 时

$$x\varphi_1(y\varphi_2 1)=x\varphi_1 y,\quad (x\varphi_1 y)\varphi_2(x\varphi_1 1)=x\varphi_1 y$$

所以等式成立。

3) $x=y=z=2$，这时等式两端都等于2，所以等式成立。这就证明了运算φ_1对运算φ_2适合分配律。同理可证φ_2对φ_1也适合分配律。

3．由于我们对二阶方阵"两得一个"的背景还很缺乏了解。所以我们只能给出以下的两个平凡的代数运算：

$$(a_{ij})\varphi_1(b_{ij})=(a_{ij}+b_{ij}),$$

$$(a_{ij})\varphi_2(b_{ij})=(a_{ij}\cdot b_{ij}).$$

显然，这两个运算都是阵的元素直接相加与直接相乘，因而它们都自然满足结合律及交换律。还有，φ_2对φ_1适合分配律。

4．φ是N的代数运算是明显的。这个运算φ不满足交换律，也不满足结合律，比如

$$1\varphi 2=1^2=1,\ 2\varphi 1=2^1=2,\quad \therefore\ 1\varphi 2\neq 2\varphi 1.$$

$$(3\varphi 1)\varphi 2=3^1\varphi 2=3^2=9,$$

$$3\varphi(1\varphi 2)=3\varphi 1^2=3\varphi 1=3^1=3$$

$$\therefore\quad (3\varphi 1)\varphi 2\neq 3\varphi(1\varphi 2).$$

5．φ_1与φ_2都是D^+的代数运算是明显的。二者都满足交换律也是显然的。但它们都不满足结合律。因为

$$(1\varphi_1 1)\varphi_1 2=\frac{1+1}{2}\varphi_1 2=1\varphi_1 2=\frac{1+2}{2}=\frac{3}{2},$$

$$1\varphi_1(1\varphi_1 2)=1\varphi_1\frac{1+2}{2}=1\varphi_1\frac{3}{2}=\frac{1+\frac{3}{2}}{2}=\frac{5}{4},$$

所以φ_1不满足结合律。

$$(2\varphi_2 2)\varphi_2 8=\sqrt{4}\varphi_2 8=2\varphi_2 8=\sqrt{16}=4,$$

$$2\varphi_2(2\varphi_2 8)=2\varphi_2\sqrt{16}=2\varphi_2 4=\sqrt{8}=2\sqrt{2}.$$

所以φ_2不满足结合律。

练　习　四

1．考虑复数集 C 的如下的子集：

$$\alpha_b=\{a+bi|\ a\in D,\ b\text{为一固定实数}\}.$$

这里子集 α_b 就是虚部系数为 b 的一切复数组成的．于是

$$\Sigma=\{\alpha_b|\ b\in D\}$$

是 C 的一个分类，在 Σ 之下，两个复数 α 与 β 分在一类里当且仅当它们的虚部相同。

如果令规则

$$\varphi:\ \alpha_b\longmapsto b$$

那么 φ 显然就是 Σ 与 D 之间的一个一一对应。

按本节定理中指出的一般原则，这个分类 Σ 所决定的等价关系 $\sim$ 为：

$\alpha\sim\beta$ 当且仅当 α 与 β 分在一类里。换句话说就是：

$\alpha\sim\beta$ 当且仅当 $\alpha-\beta$ 为实数。

2．考虑 D 的如下的子集：

$$\Delta_a=\{x\in D|\ a\leqslant x<a+1,\ a\in Z\}$$

这里子集 Δ_a 就是左闭右开的区间 $[a,\ a+1)$ 内一切实数组成的．于是

$$\Sigma=\{\Delta_a|\ a\in Z\}$$

是 D 的一个分类．在 Σ 之下，两个实数 u，v 分在一类里当且仅当 u 与 v 的整数部分相等。

如果令规则

$$\varphi:\ \Delta_a\longmapsto a$$

那么 φ 就是 Σ 与 Z 之间的一个一一对应。

这个分类 Σ 所决定的等价关系 $\sim$ 为

$u\sim v$ 当且仅当 u 与 v 分在一类里。

换句话说就是：

$u \sim v$当且仅当u与v的整数部分相等。

3．考虑 Z 的如下的子集：

$$S_{2n}=\{2n, 2n+1\}.$$

这里S_{2n}就是相邻二整数$2n$与$2n+1$组成的．于是

$$\Sigma=\{S_{2n} | n \in Z\}.$$

是 Z 的一个分类。

如果令规则

$$\varphi: S_{2n} \longrightarrow 2n.$$

那么φ就是Σ与偶数集之间的一个一一对应。

这个分类Σ所决定的等价关系～为

$a \sim b$当且仅当a与b分在一类里．换个说法就是：

$a \sim b$当且仅当$a=b$或者 a，b 是$2n$，$2n+1$ 这样两个相邻整数．

4．

1）令

$\sim_1$：$a \sim_1 b$当且仅当a与b的奇偶性相同．明显地，$\sim_1$是等价关系．于是由$\sim_1$所决定的分类为

a与b分在一类当且仅当$a \sim_1 b$．

换句话说就是：

a与b分在一类当且仅当a, b的奇偶性相同。

这样，$A=\{1,2,3,4,5,6\}$ 中的 1，3，5 分在一类里，2，4，6 分在一类里．所以

$$\Sigma=\{\{1,3,5\},\{2,4,6\}\}$$

就是$\sim_1$所决定的分类，它含有两个子集。

2）设想：根据本节定理的一般原理，我们可以首先把集合$A=\{1,2,3,4,5,6\}$ 做成一个分类Σ，使 Σ 恰好含有三个子集就行了，这当然是很容易办到的．比如令$A_1=\{1,2\}$，$A_2=\{3,4\}$，$A_3=\{5,6\}$，那么$\Sigma=\{A_1, A_2, A_3\}$就是A的分成三个子集的一个分类．于是按此分类Σ决定一个等价关系，设为$\sim_2$．从而由$\sim_2$所决定的分类恰好就是Σ．把以上想法写出来就是：

首先做成A的一个含有三个子集的分类．比如是 $\Sigma=\{\{1,2\},\{3,4\},\{5,6\}\}$．令$\Sigma$所决定的等价关系为

$\sim_2$: $a\sim_2 b$当且仅当a与b在Σ之下分在一类里．

从而由$\sim_2$所决定的分类就是当初的 Σ．这样，等价关系 $\sim_2$所决定的分类就含有三个元素．

3）既然要求分类中只有一个子集，因而就是任二元素都在一个类里．所以，这就相当于任二元素都应有关系．于是可令

$\sim_3$: $a\sim_3 b$, $\forall a$, $b\in A$.

当然，等价关系$\sim_3$所决定的分类Σ只有一个类，即$\Sigma=\{\{1,2,3,4,5,6\}\}$．

4）六个元素分成六类，就是每一个元素各成一类，于是令

$\sim_4$: $a\sim_4 b$当且仅当 $a=b$．

这样，等价关系$\sim_4$所决定的分类Σ就含有六个子集，即 $\Sigma=\{1,2,3,4,5,6\}$．

习 题 七

1．$A\cup B=\{0,1,2,3,4\}$，$A\cap B=\{2,4\}$，$A-B=\{1,3\}$，$B-A=\{0\}$，

$B\times A=\{(0,1),(0,2),(0,3),(0,4),(2,1),(2,2),(2,3),(2,4),(4,1),(4,2),(4,3),(4,4)\}$

$P(B)=\{\{0\},\{2\},\{4\},\{0,2\},\{0,4\},\{2,4\},\{0,2,4\},\phi\}$．

2．设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

规定

$$\varphi: (x_1, x_2, \cdots, x_n) \longmapsto (y_1, y_2 \cdots, y_n),$$

其中

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{cases} \qquad (*)$$

显然，规则φ是$D^{(n)}$的一个变换，使

φ: $(0, \cdots, 0, 1, 0, \cdots, 0) \longmapsto (a_{1j}, a_{2j}, \cdots, a_{nj})$, $j = 1, 2, \cdots, n$.

我们指出，φ是单满变换.

事实上，对于给定的n个数 $y_1, y_2, \cdots, y_n$, 依据克莱姆法则，有且只有一组数$x_1, x_2, \cdots, x_n$ 使（*）式成立，这就表明由（*）式所规定的变换φ是单满变换.

再有，由克莱姆法则，可以从（*）式解出各个x_i的数值. 即用 $y_1, y_2, \cdots, y_n$把 $x_1, x_2, \cdots, x_n$ 表达出来：

$x_i = \dfrac{d_i}{d}$, $d = \det A$, d_i是用 $y_1, y_2, \cdots, y_n$ 代替行列式d中第 i 列得到的n阶行列式. 于是

$$\psi: (y_1, y_2, \cdots, y_n) \longmapsto \left(\frac{d_1}{d}, \frac{d_2}{d}, \cdots, \frac{d_n}{d}\right)$$

就是变换φ的逆变换.

3. 此题有些不同的解法，关键在于我们想要说明n阶方阵与n维向量的什么联系. 目前大体上可以考虑以下三种回答方法.

1) φ_1: $A \longmapsto \{\alpha_1, \alpha_2, \cdots, \alpha_n\}$, α_i是A的第i个行向量.

2) φ_2: $A \longmapsto L(\alpha_1, \alpha_2, \cdots, \alpha_n) = \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n \mid k_i \in D\}$. 比处$L(\alpha_1, \alpha_2, \cdots, \alpha_n)$就是$\alpha_1, \alpha_2, \cdots, \alpha_n$生成的子空间.

3) φ_3: $A \longmapsto L_A$,

此处L_A表示以A为系数阵的齐次线性方程组的解空间.

显然，φ_1，φ_2，φ_3都是$M_n(D)$到$P(D^{(n)})$的映射．其中φ_1就是把A的行看做是一组n维向量；φ_2则认为以A的行向量生成一个唯一的子空间；φ_3考虑的是把A看做一个齐次线性方程组，从而它确定唯一的一个子空间—解空间．

4．首先，φ_1与φ_2都是$M_n(F)$的变换是明显的．

其次，φ_1与φ_2都不是单的也都不是满的．对此我们可以做如下解释．对于任一$(a_{ij})\in M_n(F)$，在φ_1之下，它的象为

$$B=(b_{ij}),\qquad b_{ij}=\frac{a_{ij}+a_{ji}}{2}.$$

于是B的转置阵$B'=(b'_{ij})$，而$b'_{ij}=b_{ji}=\dfrac{a_{ji}+a_{ij}}{2}=\dfrac{a_{ij}+a_{ji}}{2}=b_{ij}$.

这就说明B是对称阵．即φ_1把任何一个n阶方阵都射成为对称阵，而且明显可见，对称阵(a_{ij})在φ_1之下的象就是(a_{ij})本身，即$\varphi_1((a_{ij}))=(a_{ij})$．但是$M_n(F)$中不全是对称阵，据此可以肯定$\varphi_1$既不是满的也不是单的．

类似地，(a_{ij})在φ_2之下的象为$C=(c_{ij})$，此处$c_{ij}=\dfrac{a_{ij}-a_{ji}}{2}$．由此容易得出$C$是反对称阵：$C'=-C$．进而便可得出$\varphi_2$既不是满的，也不是单的．

最后，我们已经看到，若A是对称阵，那么A在φ_1之下的象就是A本身，即$\varphi_1(A)=A$．所以一切对称阵都在子集S_1中．另一方面，如果$\varphi_1(A)=A$．则有

$$a_{ij}=\frac{a_{ij}+a_{ji}}{2},\quad 2a_{ij}=a_{ij}+a_{ji},\quad a_{ij}=a_{ji}$$

所以A为对称阵．由此得知子集S_1恰好由一切对称阵所组成．

同理可得，子集S_2恰好由一切反对称阵所组成．

5．这里如果不做任何要求，便可随意给$A=\{a,\ b\}$规定两个“两得一个”的办法，如

1) $x\varphi_1 y=a,\forall x,y\in A$；

2) $x\varphi_2 y = b$, $\forall x, y \in A$.

显然，如此简单安排的两个代数运算满足结合律与交换律．有意思的是，不难想到，分配律是不成立的．事实上

$$x\varphi_1(y\varphi_2 z) = x\varphi_1 b = a, \quad (x\varphi_1 y)\varphi_2(x\varphi_1 z) = a\varphi_2 a = b,$$

$$\therefore \quad x\varphi_1(y\varphi_2 z) \neq (x\varphi_1 y)\varphi_2(x\varphi_1 z).$$

$$(y\varphi_2 z)\varphi_1 x = b\varphi_1 x = a, \quad (y\varphi_1 x)\varphi_2(z\varphi_1 x) = a\varphi_2 a = b,$$

$$\therefore \quad (y\varphi_2 z)\varphi_1 x \neq (y\varphi_1 x)\varphi_2(z\varphi_1 x).$$

完全类似地，可知

$$x\varphi_2(y\varphi_1 z) \neq (x\varphi_2 y)\varphi_1(x\varphi_2 z),$$

$$(y\varphi_1 z)\varphi_2 x \neq (y\varphi_2 x)\varphi_1(z\varphi_2 x).$$

如果我们把$A=\{a,b\}$的元素a与b当做代数学的对象来看的话．那么就可以凭借我们关于两个元素的集合的全部经验，赋予 $A=\{a,b\}$以这样或那样的背景．比如，我们把$A=\{a,b\}$中的a与b想象成是－1与1的代表，那么就可以象练习三第一题那样来安排$A=\{a,b\}$的运算．又如，我们认为 a,b 代表1与2两个数，那么自然可以象练习三第二题那样来安排$A=\{a,b\}$的代数运算．总而言之，我们对两个元素组成的集合有什么样的认识，都可以通过 $A=\{a,b\}$把它表达出来，从而$A=\{a,b\}$ 的运算方法也就具有了某种具体的含义．

其实，一般地，$A=\{a,b\}$ 共有16种可能的运算方法．这些运算方法可以用一定的表格形式具体的列举出来．即

①	a	b
a	a	a
b	a	a

②	a	b
a	a	a
b	a	b

③	a	b
a	a	b
b	a	a

④	a	b
a	b	a
b	a	a

⑤	a	b
a	a	a
b	b	a

⑥	a	b
a	a	a
b	b	b

⑦	a	b
a	b	b
b	a	a

⑧	a	b
a	a	b
b	a	b

⑨	a	b
a	b	a
b	b	a

⑩	a	b
a	a	b
b	b	a

⑪	a	b
a	b	a
b	a	b

⑫	a	b
a	a	b
b	b	b

⑬	a	b
a	b	a
b	b	b

⑭	a	b
a	b	b
b	a	b

⑮	a	b
a	b	b
b	b	a

⑯	a	b
a	b	b
b	b	b

此表把左侧竖行元素与上方横行元素运算的结果写在二者的交叉处，即

φ	$\cdots\ y\ \cdots$
$\vdots$	$\vdots$
x	$\cdots\ z\ \cdots$
$\vdots$	$\vdots$

其中 $x\varphi y=z$。

比如第6个运算⑥就是

$$a⑥a=a,\quad a⑥b=a,$$

$$b⑥a=b,\quad b⑥b=b.$$

如果令 a,b 分别代表1与2，运算方法一个是求大公约，一个是求小公倍．那么这两个运算就是上面列出的运算②与⑫．而把a，b当做-1与1来看时，练习三第一题给出的两个运算φ_1与φ_2就分别是这里的⑪与⑥．

6．φ是代数运算是明显的．φ满足交换律也是容易知道的．因为

$$a\varphi b=a+b-ab,$$

$$b\varphi a=b+a-ba$$

$$\therefore\quad a\varphi b=b\varphi a.$$

下面验证φ也满足结合律．

$$\begin{aligned}(a\varphi b)\varphi c&=(a+b-ab)\varphi c=a+b-ab+c-(a+b-ab)c\\&=a+b+c-(ab+bc+ca)+abc,\\a\varphi(b\varphi c)&=a\varphi(b+c-bc)=a+(b+c-bc)-a(b+c-bc)\\&=a+b+c-(ab+bc+ca)+abc.\end{aligned}$$

$$\therefore\quad (a\varphi b)\varphi c=a\varphi(b\varphi c).$$

即φ满足结合律．

7．规则$\sim_1$无法判断零多项式0与其它任何一个多项式 $h(x)$ 是有关系还是没有关系．这是因为当$f(x)=0$时，式子$\dfrac{f(x)}{(f(x),f'(x))}$

没有意义。因而无法衡量 0 与$h(x)$是否有关系。这样，规则$\sim_1$不是$D[x]$的关系，当然更谈不上是等价关系或者不是等价关系。

规则$\sim_2$对于非 0 多项式来说与$\sim_1$同效。而当$f(x)$与$g(x)$之中有 0 多项式时，等式

$$f(x)(g(x),g'(x))=g(x)(f(x),f'(x))$$

总是成立的。这样，规则$\sim_2$对$D[x]$中任意两个多项式$f(x)$与$g(x)$，都能使得

$$f(x)\sim_2 g(x),\quad f(x)\nsim_2 g(x)$$

恰有一种情形成立。所以$\sim_2$是$D[x]$的一个关系。但是$\sim_2$不是等价关系。这是因为$\sim_2$不具有传递性。比如，对任意的$f(x)$，$g(x)$，总有

$$f(x)\sim_2 0,\quad 0\sim_2 g(x).$$

但不会总有$f(x)\sim_2 g(x)$。事实上$1\nsim_2 x$。

$\sim_3$是等价关系。由它所决定的分类为

$$\Sigma=\{\{0\},\ \dot{D},\ [p_1(x)],[p_1(x)p_2(x)],\cdots,[p_1(x)\cdots p_r(x)],\cdots\}.$$

其中$[p_1(x)\cdots p_r(x)]=\{ap_1^{m_1}(x)\cdots p_r^{m_r}(x)\mid m_i\in N,\ a\in\dot{D}\}$，$p_i(x)$不可约。

8．1）给定一个规则

φ：向量集合S_1与S_2分在一类里当且仅当S_1与S_2的秩相等。

显然，规则φ是$P(D^{(m)})$的一个分类，记作Σ，于是

$$\Sigma=\{r_0,r_1,r_2,\cdots,r_n\},$$

其中$r_m=\{S\in P(D^{(m)})\mid S\text{的秩为}m\}$。

由分类Σ所决定的等价关系$\sim$为

$S_1\sim S_2$当且仅当S_1与S_2分在一类里，换句话说就是

$S_1\sim S_2$当且仅当S_1与S_2的秩相等。

2）给定一个规则$\sim$：

$S_1\sim S_2$当且仅当S_1与S_2的极大无关组等价（简称S_1与S_2等价）。显然，规则$\sim$是$P(D^{(m)})$的一个等价关系。由$\sim$所决定的分类Σ为

S_1与S_2分在一类里当且仅当$S_1 \sim S_2$。换句话说就是

S_1与S_2分在一类里当且仅当S_1与S_2等价。

应该想到：如果S_1与S_2等价，那么S_1与S_2的秩相等。反之则未必。因此，1）与2）中的分类（等价关系）并不一样。明显地，2）中的分类比1）中的分类要细得多，分成的子集远不止$n+1$个。我们可以大体上把2）中的分类表达如下：

$$\Sigma = \{\{0\}, S_1, T_1, \cdots, S_2, T_2, \cdots, S_m, T_m, \cdots\},$$

其中$S_m, T_m, \cdots$表示秩为m的等价子集做成的类，而$r_m = \{S_m, T_m, \cdots\}$.

第八章　矩阵的运算

练　习　一

1　解

1) $\begin{pmatrix} 5 & 2 & -2 \\ 4 & -1 & 3 \\ -4 & 3 & 0 \end{pmatrix}$,　　2) $\begin{pmatrix} 4 & 1 & 0 & -1 \\ 3 & 0 & 0 & 6 \\ 3 & -2 & 3 & 2 \end{pmatrix}$.

2　解

$$X = \begin{pmatrix} 1 & 4 & 2 \\ 0 & 4 & 3 \\ 0 & -2 & 2 \end{pmatrix}.$$

3　解

1) $\begin{pmatrix} 3 \\ 8 \end{pmatrix}$,　　2) $(a_1b_1 + a_2b_2 + \cdots + a_nb_n)$,

3) $\begin{pmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \cdots & \cdots & \cdots & \cdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{pmatrix}$,　　4) $(ax^2 + 2bxy + cy^2 + 2dx + 2ey + f)$.

4　解

1) $\begin{pmatrix} 2 & 2 & -2 \\ 2 & 0 & 0 \\ 4 & -4 & -2 \end{pmatrix}$,

2) $\begin{pmatrix} b-ac & a^2+b^2+c^2-b-c-ab & b^2-a^2+2ac-2c \\ c-bc & 2ac-2b & a^2+b^2+c^2-b-c-ab \\ 3-2a-c^2 & c-bc & b-ac \end{pmatrix}$.

5　解

1) $\begin{pmatrix}1 & n\\0 & 1\end{pmatrix}$，　2) $\begin{pmatrix}\cos n\theta & -\sin n\theta\\ \sin n\theta & \cos n\theta\end{pmatrix}$，　3) $\begin{pmatrix}1 & n & \frac{n(n-1)}{2}\\ 0 & 1 & n\\ 0 & 0 & 1\end{pmatrix}$.

6　解

$$AA'=\begin{pmatrix}5 & 4 & 5\\4 & 5 & 7\\5 & 7 & 10\end{pmatrix},\qquad A'A=\begin{pmatrix}14 & 7\\ 7 & 6\end{pmatrix}.$$

7　解

1) $f(A)=\begin{pmatrix}5 & 1 & 3\\8 & 0 & 3\\-2 & 1 & -2\end{pmatrix}$，　2) $f(A)=\begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}$.

8　解　将A，B分块：

$$A=\left(\begin{array}{cc|cc}\alpha & 1 & & 0\\ 0 & \alpha & & \\ \hline & & \beta & 1\\ & 0 & 1 & \beta\end{array}\right)=\begin{pmatrix}A_1 & 0\\0 & A_2\end{pmatrix},$$

$$B=\left(\begin{array}{cc|cc}\alpha & 0 & & 0\\ 1 & \alpha & & \\ \hline & & \beta & 0\\ & 0 & 1 & \beta\end{array}\right)=\begin{pmatrix}B_1 & 0\\0 & B_2\end{pmatrix}.$$

$$A^2=\begin{pmatrix}A_1 & \\ & A_2\end{pmatrix}^2=\begin{pmatrix}A_1^2 & \\ & A_2^2\end{pmatrix}=\begin{pmatrix}\alpha^2 & 2\alpha & & 0\\ 0 & \alpha^2 & & \\ & & \beta^2+1 & 2\beta\\ & 0 & 2\beta & \beta^2+1\end{pmatrix};$$

$$ABA=\begin{pmatrix}A_1 & \\ & A_2\end{pmatrix}\begin{pmatrix}B_1 & \\ & B_2\end{pmatrix}\begin{pmatrix}A_1 & \\ & A_2\end{pmatrix}$$

$$=\begin{pmatrix}A_1B_1A_1 & \\ & A_2B_2A_2\end{pmatrix}=\begin{pmatrix}\alpha^3+\alpha & 2\alpha^3+1 & & 0\\ \alpha^2 & \alpha^3+\alpha & & \\ & & \beta^3+2\beta & 2\beta^2+1\\ & 0 & 3\beta^2 & \beta^3+2\beta\end{pmatrix}.$$

$$B^3=\begin{pmatrix}B_1 & \\ & B_2\end{pmatrix}^3=\begin{pmatrix}B_1^3 & \\ & B_2^3\end{pmatrix}=\begin{pmatrix}\alpha^3 & 0 & & 0 \\ 3\alpha^2 & \alpha^3 & & \\ & & \beta^3 & 0 \\ & 0 & 3\beta^2 & \beta^3\end{pmatrix}.$$

9　证明　显然 $x=-A+B$ 是方程的解。如方程还有解，令为C，即$A+C=B$，则

$$C=-A+B.$$

这就证明了方程

$$A+X=B$$

有且只有一解。

10　证明　充分性是明显的。现在来证必要性。若$k=0$，问题得证。否则，若$k\neq 0$，则

$$\frac{1}{k}\ (kA)=\frac{1}{k}0, \qquad \text{即}\quad A=0.$$

11　解

1) $\begin{pmatrix}c & b \\ 0 & c\end{pmatrix}$，$b$，$c$为任意数；

2) $\begin{pmatrix}a & c & d \\ 0 & a & c \\ 0 & 0 & a\end{pmatrix}$，$a, c, d$为任意数。

12　解　不总成立。例如，当

$$A=\begin{pmatrix}1 & & \\ & 2 & \\ & & 3\end{pmatrix},\quad B=\begin{pmatrix}-1 & & \\ & -2 & \\ & & 3\end{pmatrix}$$

则

$$A+B=\begin{pmatrix}0 & & \\ & 0 & \\ & & 6\end{pmatrix}.$$

于是　$\det(A+B)=0$，但　$\det A+\det B=12$。

练　习　二

1　解

1）可逆，其 $A^{-1}=\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$；

2）不可逆；

3）可逆，其 $A^{-1}=\begin{pmatrix} 1 & 3 & -2 \\ -\frac{3}{2} & -3 & \frac{5}{2} \\ 1 & 1 & -1 \end{pmatrix}$；

4）可逆，其 $A^{-1}=\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 1 & 2 & -3 \\ 1 & -2 & -1 & 2 \end{pmatrix}$；

5）可逆，其 $A^{-1}=\begin{pmatrix} a_1^{-1} & & & \\ & a_2^{-1} & & \\ & & \ddots & \\ & & & a_n^{-1} \end{pmatrix}$。

2　解

1）$X=\begin{pmatrix} -17 & -28 \\ -4 & -6 \end{pmatrix}$，　　2）$X=\begin{pmatrix} 2 & 9 & -5 \\ -2 & -8 & 6 \\ -4 & -14 & 9 \end{pmatrix}$，

3）$X=\begin{pmatrix} -13 & -75 & 30 \\ 9 & 52 & -21 \\ 21 & 120 & -47 \end{pmatrix}$，　4）$X=\begin{pmatrix} a & b \\ a-2 & b-5 \\ -2a+1 & -2b+4 \end{pmatrix}$，

a,b任意。

3　1）证明　因为

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}=\begin{pmatrix} AA^{-1} & 0 \\ 0 & BB^{-1} \end{pmatrix}=\begin{pmatrix} I_n & \\ & I_n \end{pmatrix}=I_{2n},$$

$$\begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} A^{-1}A & 0 \\ 0 & B^{-1}B \end{pmatrix} = \begin{pmatrix} I_n & \\ & I_n \end{pmatrix} = I_{2n}.$$

所以 $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$为可逆的，其$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & B^{-1} \end{pmatrix}$。

同理可证：$\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}$为可逆的，其$\begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & B^{-1} \\ A^{-1} & 0 \end{pmatrix}$。

2) 解

$$\begin{pmatrix} 3 & -2 & 0 & 0 \\ 5 & -3 & 0 & 0 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -3 & 2 & 0 & 0 \\ -5 & 3 & 0 & 0 \\ 0 & 0 & -1 & 4 \\ 0 & 0 & 1 & -3 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 & 3 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ -3 & 2 & 0 & 0 & 0 \\ 2 & -1 & 0 & 0 & 0 \end{pmatrix}.$$

4 解

$$(A+2I_3)^{-1}(A^2-4I_3) = (A+2I_3)^{-1}(A+2I_3)\ (A-2I_3)$$
$$= A-2I_3$$
$$= \begin{pmatrix} -3 & 0 & 0 \\ 1 & -3 & 0 \\ 1 & 1 & -3 \end{pmatrix};$$

$$(A+2I_3)^{-1}(A-2I_3) = \begin{pmatrix} -3 & 0 & 0 \\ 4 & -3 & 0 \\ 0 & 4 & -3 \end{pmatrix}.$$

5 证明 因为

$$(I_n - A)(I_n + A + A^2 + \cdots + A^{k-1})$$
$$= I_n + A + A^2 + \cdots + A^{k-1} - A - A^2 - \cdots - A^{k-1} - A^k$$
$$= I_n - A^k$$
$$= I_n,$$

同理

$$(I_n + A + A^2 + \cdots + A^{k-1})(I_n - A) = I_n.$$

故

$$(I_n - A)^{-1} = I_n + A + A^2 + \cdots + A^{k-1}.$$

6　解　有。比如

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

则

$$AB = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

7　证明　因为

$$(-A)(-A^{-1}) = I_n, \quad (-A^{-1})(-A) = I_n.$$

故$-A$为可逆阵，且$(-A)^{-1} = -A^{-1}$。

8　证明

$$A^{-1}B = A^{-1}B(AA^{-1}) = A^{-1}(BA)A^{-1} = A^{-1}ABA^{-1} = BA^{-1}.$$

9　证明　因为

$$A^* \cdot \frac{1}{|A|}A = \frac{1}{|A|}A^*A = \frac{1}{|A|}|A|\,I_n = I_n,$$

$$\frac{1}{|A|}A \cdot A^* = \frac{1}{|A|}(AA^*) = \frac{1}{|A|}|A|\,I_n = I_n.$$

于是，A^*为可逆阵，其逆阵

$$(A^*)^{-1} = \frac{1}{|A|}A.$$

10　证明　由$A^2 = A$，得$A^2 - A = 0$，即$A(A - I_n) = 0$。若A可逆，则用A^{-1}左乘两端，得

$$A - I_n = 0, \quad 即A = I_n.$$

这与$A \neq I_n$矛盾。故A不可逆。

练　习　三

1　解

1）A的标准形为$\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$，

$$D_3\left(-\frac{1}{9}\right)P_{32}(-5)D_1(-1)P_{31}(2)C_{12}AP_{12}(1)P_{13}(4)C_{23}P_{23}(-2)$$

$$=\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}.$$

2）A的标准形为$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}$，

$$P_{42}(1)P_{32}(3)\ D_1\left(\frac{1}{2}\right)P_{41}(1)P_{31}(-2)C_{12}A$$

$$P_{12}\left(-\frac{1}{2}\right)P_{13}\left(\frac{1}{2}\right)P_{14}(-1)C_{24}P_{21}(-3)D_3\left(\frac{1}{14}\right)$$

$$=\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}.$$

2　解

1）$\frac{1}{21}\begin{pmatrix} -3 & 23 & 22 \\ 6 & -32 & -37 \\ 3 & -2 & -1 \end{pmatrix}$.　　　2）$\frac{1}{6}\begin{pmatrix} -4 & 2 & 1 & -1 \\ -4 & -10 & 7 & -1 \\ 8 & 2 & -2 & 2 \\ -2 & 4 & -1 & 1 \end{pmatrix}$.

3　解

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}\xrightarrow{P_{21}(a^{-1})}\begin{pmatrix} a & 0 \\ 1 & a^{-1} \end{pmatrix}\xrightarrow[P_{12}(1-a^{-1})]{}\begin{pmatrix} a & a-1 \\ 1 & 1 \end{pmatrix}\xrightarrow{P_{12}(1-a)}$$

$\begin{pmatrix}1&0\\1&1\end{pmatrix}$.

从而，有

$$\begin{pmatrix}1&1-a\\0&1\end{pmatrix}\begin{pmatrix}1&0\\a^{-1}&1\end{pmatrix}\begin{pmatrix}a&\\&a^{-1}\end{pmatrix}\begin{pmatrix}1&1-a^{-1}\\0&1\end{pmatrix}=\begin{pmatrix}1&0\\1&1\end{pmatrix}.$$

于是

$$\begin{pmatrix}a&0\\0&a^{-1}\end{pmatrix}=\begin{pmatrix}1&0\\a^{-1}&1\end{pmatrix}^{-1}\begin{pmatrix}1&1-a\\0&1\end{pmatrix}^{-1}\begin{pmatrix}1&0\\1&1\end{pmatrix}\begin{pmatrix}1&1-a^{-1}\\0&1\end{pmatrix}^{-1}$$

$$=\begin{pmatrix}1&0\\-a^{-1}&1\end{pmatrix}\begin{pmatrix}1&a-1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\1&1\end{pmatrix}\begin{pmatrix}1&a^{-1}-1\\0&1\end{pmatrix}.$$

4　证明　令 $\operatorname{rank}A=r$，那么存在可逆阵 P，Q，使得

$$A=P\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix}Q \quad (r\text{个}1)$$

$$=P\left(\begin{pmatrix}1&&&\\&0&&\\&&\ddots&\\&&&0\end{pmatrix}+\begin{pmatrix}0&&&\\&1&&\\&&\ddots&\\&&&0\end{pmatrix}+\cdots+\begin{pmatrix}0&&&&\\&\ddots&&&\\&&1&&\\&&&0&\\&&&&\ddots\\&&&&&0\end{pmatrix}\right)Q$$

$$=P\begin{pmatrix}1&&&\\&0&&\\&&\ddots&\\&&&0\end{pmatrix}Q+P\begin{pmatrix}0&&&\\&1&&\\&&\ddots&\\&&&0\end{pmatrix}Q+\cdots+P\begin{pmatrix}0&&&&\\&\ddots&&&\\&&1&&\\&&&0&\\&&&&\ddots\\&&&&&0\end{pmatrix}Q$$

$=B_1+B_2+\cdots+B_r$，

其中 $\operatorname{rank}B_i=1$，$i=1,2,\cdots,r$.

5　证明　由于 $\operatorname{rank}A=r$，所以只对A施行行的初等变换可

化为

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

于是存在可逆矩阵P，使

$$PA=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

从而PAP^{-1}的后$n-r$行全为0。

6　证明　由于$|A|=1$，据本节定理2的推论可知

$$\begin{aligned} A &= P_1P_2\cdots P_t\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \\ &= P_1P_2\cdots P_t, \end{aligned}$$

其中P_i为消法矩阵，$i=1,2,\cdots,t$。

7　证明　由于$AA^{-1}=I_n$，所以有

$$\det AA^{-1}=\det I_n=1.$$

据本节定理4，有

$$\det A\cdot\det A^{-1}=1.$$

故

$$\det A^{-1}=(\det A)^{-1}.$$

8　解　由于$AB=I_n$，所以有

$$\det AB=\det I_n=1.$$

据本节定理4，有

$$\det A \cdot \det B = 1.$$

因此

$$\det A \neq 0, \ \det B \neq 0.$$

于是，A，B都是可逆的，且$A^{-1}=B$，$B^{-1}=A$。即A与B互逆。

9　证明　因为由题设知

$$cI_n = -aA^2 - bA,$$

即

$$I_n = -\frac{a}{c}A^2 - \frac{b}{c}A = A\left(-\frac{a}{c}A - \frac{b}{c}\right).$$

据上题可知，A为可逆矩阵，其 $A^{-1} = -\frac{a}{c}A - \frac{b}{c}$.

10　证明　由本节定理6，有

$$\operatorname{rank}C \leqslant \min(\operatorname{rank}A, \ \operatorname{rank}B).$$

由于$m>n$，所以有

$$\operatorname{rank}A \leqslant n, \ \operatorname{rank}B \leqslant n.$$

于是

$$\operatorname{rank}C \leqslant n.$$

又因为$C=AB$是m阶方阵，$n<m$，故

$$\operatorname{rank}C < m.$$

11　解　不一定。比如

$$A=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}, \ B=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}, \ C=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}.$$

则$ACB=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$，其秩为1，而$AB=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$，其秩为0。

12　证明

1）因$\operatorname{rank}A=n$，所以$m\geqslant n$。不妨设A的前n行线性无关，且构成n阶可逆阵设为A_1。A的后$m-n$行构成的矩阵设为A_2。则

$$AB=\begin{pmatrix}A_1\\ A_2\end{pmatrix}B=\begin{pmatrix}A_1B\\ A_2B\end{pmatrix}.$$

于是

$$\text{rank}\,AB \geqslant \text{rank}\,A_1B = \text{rank}\,B.$$

但据本节定理6，有

$$\text{rank}\,AB \leqslant \text{rank}\,B.$$

故

$$\text{rank}\,AB = \text{rank}\,B.$$

2) 同理可证.

习 题 八

1 证明 设

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix}$$

与A可交换，于是

$$AX = \begin{pmatrix} a_1x_{11} & a_1x_{12} & \cdots & a_1x_{1n} \\ a_2x_{21} & a_2x_{22} & \cdots & a_2x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_nx_{n1} & a_nx_{n2} & \cdots & a_nx_{nn} \end{pmatrix}$$

$$XA = \begin{pmatrix} a_1x_{11} & a_2x_{12} & \cdots & a_nx_{1n} \\ a_1x_{21} & a_2x_{22} & \cdots & a_nx_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_1x_{n1} & a_2x_{n2} & \cdots & a_nx_{nn} \end{pmatrix}.$$

由$AX = XA$，得

$$a_ix_{ij} = a_jx_{ij}, \quad (i,j = 1,2,\cdots,n).$$

亦即

$$a_ix_{ij} - a_jx_{ij} = 0, \quad (a_i - a_j)x_{ij} = 0.$$

由于当$i \neq j$时，$a_i \neq a_j$. 故由上式便有

$$x_{ij} = 0, \quad (i \neq j, \ i, \ j = 1,2,\cdots, \ n).$$

于是X只能是对角形阵.

2　证明　设

$$X=\begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1r} \\ X_{21} & X_{22} & \cdots & X_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ X_{r1} & X_{r2} & \cdots & X_{rr} \end{pmatrix}$$

与A可交换（其中X_{ij}是$n_i\times n_j$矩阵），则由$AX=XA$，可得

$$(a_iI_i)X_{ij}=X_{ij}(a_jI_j),\quad (i,j=1,2,\cdots,r).$$

亦即

$$a_iX_{ij}=a_jX_{ij},\quad (a_i-a_j)X_{ij}=0.$$

由于当$i\neq j$时，$a_i\neq a_j$，故由上式便有

$$X_{ij}=0,\quad (i\neq j,\ i,j=1,2,\cdots,r).$$

于是X只能是准对角形阵。

3　证明

1）因为

$$AE_{12}=\begin{pmatrix} 0 & a_{11} & 0 & \cdots & 0 \\ 0 & a_{21} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & a_{n1} & 0 & \cdots & 0 \end{pmatrix}$$

$$E_{12}A=\begin{pmatrix} a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

且$AE_{12}=E_{12}A$可知，当$k\neq 1$时，$a_{k1}=0$，当$k\neq 2$时$a_{2k}=0$；

2）因为

$$AE_{ij}=\begin{pmatrix} 0 & \cdots & 0 & a_{1i} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & a_{2i} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & a_{ni} & 0 & \cdots & 0 \end{pmatrix}$$

（j列）

$$E_{ij}A=\begin{pmatrix} 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \begin{matrix} \\ \\ \\ (i行) \\ \\ \\ \\ \end{matrix},$$

且$AE_{ij}=E_{ij}A$可知，当$k\neq i$时 $a_{ki}=0$，当$k\neq j$时$a_{jk}=0$，且$a_{ii}=a_{jj}$；

3）因为A与所有n阶矩阵可交换，所以A必与 E_{ij} 可交换，即$AE_{ij}=E_{ij}A$。由 2）知

$$a_{ij}=0,\quad (i\neq j),且\ a_{ii}=a_{jj},\quad (i,j=1,2,\cdots,n)。$$

故A为纯量阵，即 $A=aI_n$。

4　证明

$$A(B_1+B_2)=AB_1+AB_2=B_1A+B_2A=(B_1+B_2)A;$$

$$A(B_1B_2)=(AB_1)B_2=B_1(AB_2)=(B_1B_2)A;$$

$$AB_1{}^k=(AB_1)B_1{}^{k-1}=B_1(AB_1{}^{k-1})=\cdots=B_1{}^kA。$$

5　证明　设

$$f(\lambda)=a_0\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n,$$

则

$$f(A)=a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI_n。$$

而A与B可交换，便有

$$\begin{aligned} f(A)B&=(a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI_n)B \\ &=a_0A^nB+a_1A^{n-1}B+\cdots+a_{n-1}AB+a_nI_nB \\ &=a_0BA^n+a_1BA^{n-1}+\cdots+a_{n-1}BA+a_nBI_n \\ &=B(a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI_n) \\ &=Bf(A)。\end{aligned}$$

6　解　当$AB=BA$时，成立。

7　证明　设

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}。$$

由于 $A\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$, $A\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$, $\cdots$, $A\begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}=\begin{pmatrix}0\\\vdots\\0\\0\end{pmatrix}$

可知，$a_{ij}=0$，$(i,j=1,2,\cdots,n)$，故$A=0$。

8 解

1) 两个上三角形阵之积仍为上三角形阵。事实上，设 $A=(a_{ij})$，$B=(b_{ij})$且当 $i>j$时$a_{ij}=0$，$b_{ij}=0$。令$AB=C=(c_{ij})$，只须证：当$i>j$时 $c_{ij}=0$ 即可。由于

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ij}b_{jj}+a_{ij+1}b_{j+1j}+\cdots+a_{in}b_{nj},$$

当$i>j$,时，上式的前 j 项的 $a_{it}=0$，$(t=1,2,\cdots,j)$，而后$n-j$ 项的 $b_{kj}=0$，$(k=j+1,\cdots,n)$。于是当$i>j$时，$c_{ij}=0$。

2）可逆的上三角形阵A的逆阵A^{-1}仍为上三角形。

事实上，令$A^{-1}=(b_{pq})$，则

$$b_{pq}=\frac{1}{|A|}A_{qp}.$$

由于A是上三角形阵，所以 $A_{ij}=0$ $(i<j)$。于是当$p>q$时

$$b_{pq}=\frac{1}{|A|}\ A_{qp}=\frac{1}{|A|}\cdot 0=0.$$

故$A^{-1}=(b_{pq})$为上三角形阵。

9 证明

1）由于 $\operatorname{rank}A=1$，所以A的任意两行、两列都成比例。于是不妨设其余各行的元素都是第一行的元素的倍数，即

$$A=\begin{pmatrix}a_1b_1 & a_1b_2 & \cdots & a_1b_n\\ a_2b_1 & a_2b_2 & \cdots & a_2b_n\\ \cdots & \cdots & \cdots & \cdots\\ a_nb_1 & a_nb_2 & \cdots & a_nb_n\end{pmatrix}=\begin{pmatrix}a_1\\a_2\\\cdots\\a_n\end{pmatrix}(b_1b_2\cdots b_n),$$

其中$a_1=1$；

2) $A^2=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1b_2\cdots b_n)\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1b_2\cdots b_n)$

$$=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\left(\sum_{i=1}^{n}a_ib_i\right)(b_1b_2\cdots b_n)$$

$$=k\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(b_1b_2\cdots b_n)$$

$$=kA,\ 其中k=\sum_{i=1}^{n}a_ib_i.$$

10 证明 据本章例题选解中例 7 知， $|A^*|=|A|^{n-1}$. 于是

1）当 $\operatorname{rank}A=n$ 时，则 $|A|\neq0$，所以 $|A^*|\neq0$，故$\operatorname{rank}A^*=n$；

2）当 $\operatorname{rank}A=n-1$ 时，则A中至少有一个$n-1$阶子式不为 0，因此 $\operatorname{rank}A^*\geqslant1$.

又因

$$AA^*=\begin{pmatrix}|A|&&\\&\ddots&\\&&|A|\end{pmatrix}=0.$$

据本章例题选解中的例题 6 知

$$\operatorname{rank}A+\operatorname{rank}A^*\leqslant n.$$

于是

$$\operatorname{rank}A^*\leqslant n-\operatorname{rank}A=1.$$

故

$$\operatorname{rank}A^*=1;$$

3）当$\operatorname{rank}A<n-1$时，则$A_{ij}=0$，$(i,j=1,2\cdots,n)$，所以$A^*=0$，故$\operatorname{rank}A^*=0$.

11 证明 当 $|A|=0$时，由上题知$\operatorname{rank}A^*\leqslant1$，所以 $(A^*)^*=0$，这时当然有

$$(A^*)^*=|A|^{n-2}A$$

成立；

当 $|A|\neq0$ 时，由于

$$A^*(A^*)^* = |A^*| I_n = |A|^{n-1} I_n,$$

两边都左乘以$\frac{1}{|A|}A$，得

$$\frac{1}{|A|} A A^* (A^*)^* = |A|^{n-2} A,$$

即

$$(A^*)^* = |A|^{n-2} A.$$

12　证明

1）由于 $BC=0$，所以$C'B'=0$．故可以把B的各行（B'的各列）看作是以C'为系数阵的齐次线性方程组的解．该方程组是r个未知数n个方程的齐次线性方程组．因$\operatorname{rank}C' = \operatorname{rank}C = r$，所以该齐次线性方程组只有零解，从而$B$的各行都为0，故$B=0$；

2）因$BC=C$，则 $BC-C=0$，即$(B-I_r)C=0$．由1）知，$B-I_r=0$，故$B=I_r$．

13　证明　由于$A^2=I_n$，所以$\operatorname{rank}A=n$．因为两个矩阵和的秩不大于各个矩阵的秩之和（第六章§4习题），于是有

$$\operatorname{rank}(A+I_n+A-I_n) \leqslant \operatorname{rank}(A+I_n) + \operatorname{rank}(A-I_n).$$

而

$$\operatorname{rank}(A+I_n+A-I_n) = \operatorname{rank}(2A) = n.$$

故

$$\operatorname{rank}(A+I_n) + \operatorname{rank}(A-I_n) \geqslant n.$$

又由于$A^2=I_n$，有

$$(A+I_n)(A-I_n) = A^2 - I_n = 0.$$

据本章例题选解中例6可知

$$\operatorname{rank}(A+I_n) + \operatorname{rank}(A-I_n) \leqslant n.$$

从而，有

$$\operatorname{rank}(A+I_n) + \operatorname{rank}(A-I_n) = n.$$

14　证明　由$\operatorname{rank}(A-I_n) = \operatorname{rank}(I_n-A)$，所以

$$\operatorname{rank}A + \operatorname{rank}(A-I_n) = \operatorname{rank}A + \operatorname{rank}(I_n-A).$$

而

$$\operatorname{rank}A+\operatorname{rank}(I_n-A)\geqslant\operatorname{rank}(A+I_n-A)=\operatorname{rank}I_n=n,$$

故

$$\operatorname{rank}A+\operatorname{rank}(A-I_n)\geqslant n.$$

又由于 $A^2=A$，所以有

$$A(A-I_n)=A^2-A=0.$$

据本章例题选解中例 6 可知

$$\operatorname{rank}A+\operatorname{rank}(A-I_n)\leqslant n.$$

因此

$$\operatorname{rank}A+\operatorname{rank}(A-I_n)=n.$$

第九章　对称阵在相合之下的标准形

练　习　一

1　解　$f(x_1, x_2, x_3) = (x_1 x_2 x_3) A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

$$= x_1^2 + 4x_1x_2 - 3x_1x_3 + 2x_2x_3 - x_3^2;$$

$$f(x_1, x_2, x_3, x_4) = (x_1x_2x_3x_4)\ B \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

$$= 2x_1x_2 + 4x_1x_3 - 2x_1x_4 + 6x_2x_3 + 2x_2x_4 - x_3x_4.$$

2　解

1)　$f(x_1x_2x_3) = (x_1x_2x_3) A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$

$$A = \begin{pmatrix} 2 & -\frac{3}{2} & 1 \\ -\frac{3}{2} & 2 & -\frac{5}{2} \\ 1 & -\frac{5}{2} & 0 \end{pmatrix};$$

2)　$f(x_1, x_2, x_3, x_4) = (x_1x_2x_3x_4)\ A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$

$$A=\begin{pmatrix} 0 & \frac{3}{2} & 2 & \frac{5}{2} \\ \frac{3}{2} & 0 & -1 & -1 \\ 2 & -1 & 0 & -\frac{3}{2} \\ \frac{5}{2} & -1 & -\frac{3}{2} & 0 \end{pmatrix}.$$

3　证明　因为

$$(A+B)'=A'+B'=A+B,\ (kA)'=kA'=kA.$$

所以 $A+B$，kA为对称阵。

4　解

$$\begin{pmatrix} 1 & 2 \\ 2 & 6 \end{pmatrix},\ \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix},\ \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & 0 \end{pmatrix}.$$

练　习　二

1．解①

$$1)\ \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{21}(1)]{P_{12}(1)} \begin{pmatrix} 1 & \frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{12}\left(-\frac{1}{2}\right),\ P_{13}(-1)]{P_{21}\left(-\frac{1}{2}\right),\ P_{31}(-1)}$$

---

①该题只给出用初等变换法作的解答，配方法略，以下均如此。

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \\ 0 & 0 & -1 \\ 1 & -\frac{1}{2} & -1 \\ 1 & -\frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

故把 1) 化为对角形：

$$\begin{pmatrix} 1 & & \\ & -\frac{1}{4} & \\ & & -1 \end{pmatrix},$$

其演化阵

$$P = \begin{pmatrix} 1 & -\frac{1}{2} & -1 \\ 1 & -\frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

2) $$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 1 & 1 & -1 & 0 \\ 0 & -1 & 1 & 1 \\ -1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{12}(-1),\, P_{14}(1)]{P_{21}(-1),\quad P_{41}(1)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & -1 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{32}(1)]{P_{23}(1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 1 & -1 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{24}(2)]{P_{42}(2)}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 \\ 1 & -1 & -1 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{34}(-1)]{P_{43}(-1)} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 3 \\ 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

故化2）为对角形：

$$\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 3 \end{pmatrix},$$

其演化阵

$$P = \begin{pmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

2 解

1) $$\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{12}(-1)\ \ P_{13}(1)]{P_{21}(-1),\ \ P_{31}(1)} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & -2 \\ 0 & -2 & -7 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[P_{23}\left(\frac{2}{3}\right)]{P_{32}\left(\frac{2}{3}\right)} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -\frac{25}{3} \\ 1 & -1 & \frac{1}{3} \\ 0 & 1 & \frac{5}{3} \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[D_1\left(\frac{1}{\sqrt{2}}\right),\ D_2\left(\frac{1}{\sqrt{3}}\right),\ D_3\left(\sqrt{\frac{3}{25}}\right)]{D_1\left(\frac{1}{\sqrt{2}}\right),\ D_2\left(\frac{1}{\sqrt{3}}\right),\ D_3\left(\sqrt{\frac{3}{25}}\right)}$$

$$
\begin{pmatrix}
1 & & \\
 & 1 & \\
 & & -1 \\
\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{\sqrt{3}}{15} \\
0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{3}}{3} \\
0 & 0 & \frac{\sqrt{3}}{5}
\end{pmatrix}.
$$

故1）的标准形为
$$
\begin{pmatrix}
1 & & \\
 & 1 & \\
 & & -1
\end{pmatrix},
$$

其演化阵
$$
P = \begin{pmatrix}
\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{\sqrt{3}}{15} \\
0 & \frac{1}{\sqrt{3}} & \frac{\sqrt{3}}{3} \\
0 & 0 & \frac{\sqrt{3}}{5}
\end{pmatrix}.
$$

2）
$$
\begin{pmatrix}
2 & 1 & -1 & -1 \\
1 & 0 & -2 & -2 \\
-1 & -2 & 1 & 1 \\
-1 & -2 & 1 & 1 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\xrightarrow[P_{12}\left(-\frac{1}{2}\right),\ P_{13}\left(\frac{1}{2}\right),\ P_{14}\left(\frac{1}{2}\right)]{P_{21}\left(-\frac{1}{2}\right),\ P_{31}\left(\frac{1}{2}\right),\ P_{41}\left(\frac{1}{2}\right)}
$$

$$
\begin{pmatrix}
2 & 0 & 0 & 0 \\
0 & -\frac{1}{2} & -\frac{3}{2} & -\frac{3}{2} \\
0 & -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\
0 & -\frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\
1 & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\xrightarrow[P_{23}(-3),\ P_{24}(-3)]{P_{32}(-3),\ P_{42}(-3)}
$$

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{10}{2} & \frac{10}{2} \\ 0 & 0 & \frac{10}{2} & \frac{10}{2} \\ 1 & -\frac{1}{2} & 2 & 2 \\ 0 & 1 & -3 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{34}(-1)]{P_{43}(-1)} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{10}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -\frac{1}{2} & 2 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[D_1\left(\frac{1}{\sqrt{2}}\right),\ D_2(\sqrt{2}),\ D_3\left(\sqrt{\frac{2}{10}}\right)]{D_1\left(\frac{1}{\sqrt{2}}\right),\ D_2(\sqrt{2}),\ D_3\left(\sqrt{\frac{2}{10}}\right)} \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & 0 \\ \frac{1}{\sqrt{2}} & -\frac{\sqrt{2}}{2} & \frac{2}{\sqrt{5}} & 0 \\ 0 & \sqrt{2} & -\frac{3\sqrt{2}}{\sqrt{10}} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{10}} & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$\xrightarrow[C_{23}]{C_{23}} \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 0 \\ \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{5}} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & -\frac{3\sqrt{2}}{\sqrt{10}} & \sqrt{2} & 0 \\ 0 & \frac{\sqrt{2}}{\sqrt{10}} & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

故2）的标准形为 $\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 0 \end{pmatrix}$，

其演化阵

$$P=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{5}} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & -\frac{3}{\sqrt{5}} & \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

3　解

1）二次型 $f(x_1,x_2,x_3)$ 的表示矩阵为

$$A=\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 4 \end{pmatrix}.$$

而

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 2 \\ 0 & 2 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{12}(-1)]{P_{21}(-1)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{23}(-2)]{P_{32}(-2)} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

于是

$$f(x_1,x_2,x_3)=g(y_1,y_2,y_3)=y_1^2+y_2^2.$$

其变数变换为

$$\begin{pmatrix} x_1 \\ x_2 \\ x^3 \end{pmatrix}=\begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

2）二次型 $f(x_1,x_2,x_3,x_4)$ 的表示矩阵

$$A=\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

而

$$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[P_{21}(1)]{P_{12}(1)} \begin{pmatrix} 1 & \frac{1}{2} & 1 & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[P_{12}\left(-\frac{1}{2}\right),\ P_{13}(-1),\ P_{14}(-1)]{P_{21}\left(-\frac{1}{2}\right),\ P_{31}(-1),\ P_{41}(-1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & 0 & 0 \\ 0 & 0 & -1 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & -1 \\ 1 & -\frac{1}{2} & -1 & -1 \\ 1 & \frac{1}{2} & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[P_{34}\left(-\frac{1}{2}\right)]{P_{43}\left(-\frac{1}{2}\right)} \begin{pmatrix} 1 & & & \\ & -\frac{1}{4} & & \\ & & -1 & \\ & & & -\frac{3}{4} \\ 1 & -\frac{1}{2} & -1 & -\frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

于是

$$f(x_1, x_2, x_3, x_4) = g(y_1, y_2, y_3, y_4) = y_1^2 - \frac{1}{4}y_2^2 - y_3^2 - \frac{3}{4}y_4^2.$$

其变数变换为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{1}{2} & -1 & -\frac{1}{2} \\ 1 & \frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}.$$

4　解

1）二次型的表示矩阵为

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{pmatrix}.$$

而A的标准形为

$$\begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix},$$

其相合演化阵

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{\sqrt{2}}{2} & -\frac{3}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{\sqrt{2}}{2} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{6}} \end{pmatrix}$$

于是二次型的标准形为

$$y_1^2 - y_2^2 - y_3^2.$$

其变数变换

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{\sqrt{2}}{2} & -\frac{3}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{\sqrt{2}}{2} & -\frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

2）二次型的表示矩阵为

$$A=\begin{pmatrix}1&1&0&0\\1&1&1&0\\0&1&1&1\\0&0&1&1\end{pmatrix}.$$

而A的标准形为

$$\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&-1\end{pmatrix},$$

相合演化阵

$$P=\begin{pmatrix}1&1&0&-1\\0&-1&0&1\\0&-2&1&1\\0&1&0&0\end{pmatrix}.$$

于是二次型的标准形为 $y_1^2+y_2^2+y_3^2-y_4^2$，其变数变换为

$$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}1&1&0&-1\\0&-1&0&1\\0&-2&1&1\\0&1&0&0\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}.$$

5　证明　设对称阵A的秩为r，则存在可逆阵P，使

$$P'AP=\begin{pmatrix}a_1&&&&\\&\ddots&&&\\&&a_r&&\\&&&0&\\&&&&\ddots\\&&&&&0\end{pmatrix},\quad a_i\neq 0.$$

于是

$$A=P^{-1'}\begin{pmatrix}a_1&&&&\\&\ddots&&&\\&&a_r&&\\&&&0&\\&&&&\ddots\\&&&&&0\end{pmatrix}P^{-1}$$

$$= P^{-1\prime}\left[\begin{pmatrix} a_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & a_r & & \\ & & & & 0 & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}\right] P^{-1}$$

$$= P^{-1\prime}\begin{pmatrix} a_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{pmatrix} P^{-1} + \cdots + P^{-1\prime}\begin{pmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & a_r & & \\ & & & & 0 & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix} P^{-1}$$

$$= A_1 + \cdots + A_r.$$

其中

$$A_i = P^{-1\prime}\begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & a_i & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix} P^{-1} \quad (i = 1, 2, \cdots, r)$$

即A_i是秩为 1 的对称阵。

练　习　三

1　解

秩为 0 ：只有 1 类；

秩为 1 ：正惯标为 1 或负惯标为 1 ，共 2 类；

秩为 2 ：正惯标为 2 ， 负惯标为 0 ⎫
　⋮　　正惯标为 1 ， 负惯标为 1 ⎬ 共 3 类；
　⋮　　正惯标为 0 ， 负惯标为 2 ⎭

秩为n：正惯标为n，　负惯标为 0 ⎫
　　　正惯标为 $n-1$，负惯标为 1 ⎪
　　　　　⋮　　　　　　⋮　　 ⎬ 共$n+1$类。
　　　正惯标为 0 ，　负惯标为 n ⎭

于是共分 $1+2+\cdots+n+n+1=\dfrac{(n+1)(n+2)}{2}$类。

2　证明　因为 $r=p+q$，而 $r-s=p+q-(p-q)=2q$. 故r与s的奇偶性相同。

又因为 $0\leqslant p\leqslant r, 0\leqslant q\leqslant r$，所以 $0\leqslant|p-q|\leqslant r$，即 $0\leqslant|s|\leqslant r$。

3　证明　因实对称阵A，必存在可逆阵P，使 $P'AP$ 为标准形。又由 $|A|<0$，必有

$$|P'AP|=|P'|\,|A|\,|P|=|P|^2|A|<0.$$

故 $P'AP$ 对角线上只能是 1 或 -1，而且 -1 的个数必为奇数。于是A的负惯标为正数。

4　证明　由题设知，对实二次型$f(x_1,\cdots,x_n)$的表示矩阵A，存在可逆阵Q，使

$$Q'AQ=\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}.$$

（其中 1 有 p个，-1 有 q个）

当令

$$R=\begin{pmatrix} \sqrt{a_1} & & & & & & & \\ & \ddots & & & & & & \\ & & \sqrt{a_p} & & & & & \\ & & & \sqrt{-b_1} & & & & \\ & & & & \ddots & & & \\ & & & & & \sqrt{-b_q} & & \\ & & & & & & 1 & \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix}$$

时，

$$R'Q'AQR=\begin{pmatrix} a_1 & & & & & & & \\ & \ddots & & & & & & \\ & & a_p & & & & & \\ & & & b_1 & & & & \\ & & & & \ddots & & & \\ & & & & & b_q & & \\ & & & & & & 0 & \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}.$$

于是经可逆变数变换

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = QR\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

使二次型 $f(x_1,\cdots,x_n)$ 变为

$$f(x_1,\cdots,x_n)=a_1y_1^2+\cdots+a_py_p^2+b_1y^2_{p+1}+\cdots+b_qy^2_{p+q}.$$

练　习　四

1　解

1）是；2）不是；3）不是；4）是。

2　解

1）是；2）是；3）不是。

3　解　不一定。比如

$$A=\begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}.$$

这时，对 $x=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$，　其中 $x_1\neq 0$，$x_2\neq 0$，$x_3\neq 0$，均有

$$X'AX=x_1^2+x_2^2>0.$$

但A不是正定的，而是半正定的。

4　证明　因为 $M'^{-1}(M'AM)M^{-1}=A$，故A与$M'AM$ 相合，于是$M'AM$也是对称的。据例题选解中例3，相合两对称阵标准形

相同，即$M'AM$与A的标准形相同，故$M'AM$与A的正定性一致.

5　证明　显然$A'A$为对称阵。据本章§4命题3知，$A'A$为正定的。

6　证明　因为对任一非零的$n\times 1$矩阵

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

有 $X'AX>0$，$X'BX>0$，所以必有

$$X'(A+B)X=X'AX+X'BX>0。$$

故$A+B$为正定的。

7　证明　由$|A|<0$可知，A为可逆的，A的负惯标为正数.于是存在可逆阵P，使

$$P'AP=\begin{pmatrix}1&&&&\\&\ddots&&&\\&&1&&\\&&&-1&&\\&&&&\ddots&\\&&&&&-1\end{pmatrix}$$

其中-1的个数为一正数。故取 $y_1=0,\cdots,\ y_{n-1}=0,\ y_n=1$ 则通过变数变换

$$\begin{pmatrix}x_1\\ \vdots\\ x_{n-1}\\ x_n\end{pmatrix}=P\begin{pmatrix}y_1\\ \vdots\\ y_{n-1}\\ y_n\end{pmatrix}$$

所得到的非零X：

$$X=\begin{pmatrix}x_1\\ \vdots\\ x_{n-1}\\ x_n\end{pmatrix}$$

有

$$X'AX<0。$$

习 题 九

1 证明 必要性. $(AB)'=B'A'=BA$，同时又有 $(AB)'=AB$，故$AB=BA$。

充分性. 因 $(AB)'=B'A'=BA=AB$，故AB为对称阵。

2 证明 设 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$。而$f(A)=a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_0I$。因

$$\begin{aligned}(f(A))'&=(a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI)'\\&=a_0A'^n+a_1A'^{n-1}+\cdots+a_{n-1}A'+a_nI\\&=a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI\\&=f(A).\end{aligned}$$

故 $f(A)$为对称阵。

3 证明 因为 $A^{-1}=\dfrac{1}{|A|}A^*$。由于A为正定的，故 $|A|>0$，且A^{-1}也是正定的。于是 $A^*=|A|\,A^{-1}$，因此，A^* 也为正定的。

4 证明 反证，假设对某个i，$a_{ii}\leqslant0$。于是对A施行成套的初等变换变为B，使B中（1，1）位置为a_{ii}。故B中有一阶子式 $a_{ii}\leqslant0$，所以B不是正定的。由A与B相合，因此，A也不是正定的。

5 证明 因$A=(a_{ij})$为正定的，所以A的 k 阶顺序主子式

$$|A_k|=\begin{vmatrix}a_{11}&\cdots&a_{1k}\\\cdots&\cdots&\cdots\\a_{k1}&\cdots&a_{kk}\end{vmatrix}>0,\quad k=1,2,\cdots,n.$$

而

$$|B_k|=\begin{vmatrix}a_{11}b_1^2&\cdots&a_{1k}b_1b_k\\\cdots&\cdots&\cdots\\a_{k1}b_kb_1&\cdots&a_{kk}b_k^2\end{vmatrix}=b_1^2\cdots b_k^2\begin{vmatrix}a_{11}&\cdots&a_{1k}\\\cdots&\cdots&\cdots\\a_{k1}&\cdots&a_{kk}\end{vmatrix}>0,$$

$k=1,2,\cdots n$。故B的一切顺序主子式均大于0，因此 B 为正定的。

6 证明 tI_n+A的k阶顺序主子式

$$\Delta_k = \begin{vmatrix} t + a_{11} & \cdots & a_{1k} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{k1} & \cdots & t + a_{kk} \end{vmatrix} = t^k + \cdots + |A_k|$$

是首系数为 1 的k次多项式。当t充分大后，Δ_k必大于零。于是对各阶顺序主子式$\Delta_1, \Delta_2, \cdots, \Delta_n$都有一个充分大的 t，使该阶顺序主子式大于零。因此，可以找到一个共同的充分大的t_0，使$t_0 I_n + A$ 的各阶顺序主子式均大于零。从而，对任何 $t \geqslant t_0$，$tI_n + A$ 为正定的。

7　证明　因为

$$\begin{aligned} X'AX &= n\Sigma x_i^2 - (\Sigma x_i)^2 = n\Sigma x_i^2 - (\Sigma x_i^2 + 2\sum_{1 \leqslant i < j \leqslant n} x_i x_j) \\ &= (n-1)\Sigma x_i^2 - 2\sum_{1 \leqslant i < j \leqslant n} x_i x_j \end{aligned}$$

$$= \sum_{1 \leqslant i < j \leqslant n} (x_i - x_j)^2.$$

于是对任一非零的$n \times 1$矩阵

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

恒有

$$X'AX \geqslant 0.$$

故A为半正定的。

8　解　二次型的表示矩阵

$$A = \begin{pmatrix} 1 & t & 5 \\ t & 4 & 3 \\ 5 & 3 & 1 \end{pmatrix}$$

的顺序主子式为

$$|1| = 1, \quad \begin{vmatrix} 1 & t \\ t & 4 \end{vmatrix} = 4 - t^2 = (2+t)(2-t), \quad |A| = -t^2 + 30t - 105.$$

若　$(2+t)(2-t) > 0$，则 $-2 < t < 2$；　　(1)

若　$|A| > 0$，则　$15 - 2\sqrt{30} < t < 15 + 2\sqrt{30}$.　　(2)

因此，没有共同的 t 同时满足（1）与（2）。故不管 t 取任何值，二次型 $f(x_1, x_2, x_3)$ 都不是正定的。

9　证明　设A的秩为 r，只须证$r = 0$即可。因A 为对称阵，

必存在可逆阵R，使

$$R'AR=\begin{pmatrix} I_p & & \\ & -I_q & \\ & & 0 \end{pmatrix}.$$

于是令

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}=R\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \tag{3}$$

则有

$$f(x_1,\cdots,x_n)=X'AX=(y_1\cdots y_n)R'AR\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_{p+q}^2. \tag{4}$$

由（4）可知，若 $r=p+q\neq0$，不妨设 $p>0$。于是取 $y_1=1, y_2=0,\cdots,y_n=0$，由（3）必有一个非零的$n\times1$矩阵 X，使$X'AX\neq0$。这与题设矛盾，故 $r=p+q=0$，即 $A=0$。

10 证明 由上题可知，若$r=p+q<n$时，取$y_1=0,\cdots,y_{n-1}=0$ $y_n=1$，由（3）必有一个非零的$n\times1$矩阵X，使 $X'AX=0$。这与题设矛盾，故$r=n$，即A为满秩的。

11 证明 由9题知，二次型 $f(x_1,\cdots,x_n)=X'AX$ 经变数变换（3）可化为（4）。这时由题设知，（4）中的p,q都不为0，于是取 $y_1=1$，$y_2=0,\cdots,y_p=0$，$y_{p+1}=1$，$y_{p+2}=0,\cdots,y_n=0$。从而由（3）可取到非零的$n\times1$矩阵X_0，使$X_0'AX_0=0$。

第十章　方阵在相似之下的标准形

练　习　一

1　证明　因$A\sim B$，所以存在可逆阵P，使$B=P^{-1}AP$，于是$B'=P'A'(P^{-1})'=P'A'P'^{-1}$，故$A'\sim B'$。

2　证明　$(P^{-1}AP)(P^{-1}BP)=P^{-1}APP^{-1}BP=P^{-1}ABP=$
$P^{-1}BAP=(P^{-1}BP)(P^{-1}AP)$。

3　证明　$BA=(A^{-1}A)(BA)=A^{-1}(AB)A$，故$AB\sim BA$。

4　证明　由题设知，存在可逆阵 P_1,P_2，使$C=P_1^{-1}AP_1$，$D=P_2^{-1}BP_2$。于是

$$\begin{pmatrix} C & \\ & D\end{pmatrix}=\begin{pmatrix} P_1^{-1}AP_1 & \\ & P_2^{-1}BP_2\end{pmatrix}=\begin{pmatrix} P_1^{-1} & \\ & P_2^{-1}\end{pmatrix}\begin{pmatrix} A & \\ & B\end{pmatrix}\begin{pmatrix} P_1 & \\ & P_2\end{pmatrix}$$

$$=\begin{pmatrix} P_1 & \\ & P_2\end{pmatrix}^{-1}\begin{pmatrix} A & \\ & B\end{pmatrix}\begin{pmatrix} P_1 & \\ & P_2\end{pmatrix},$$

故

$$\begin{pmatrix} A & \\ & B\end{pmatrix}\sim\begin{pmatrix} C & \\ & D\end{pmatrix}.$$

5　证明　由题设知，$a_1,a_2,\cdots,a_n$经过若干次对换后变为 a_{i_1}，a_{i_2}，$\cdots$，a_{i_n} 且$a_{i_k}=b_k$，$k=1,2,\cdots,n$。这种对换用在对角形阵上就是成套的换法变换，即

$$B=T_s^{-1}\cdots T_1^{-1}AT_1\cdots T_s=T^{-1}AT,$$

其中T_i为换法阵，$T=T_1\cdots T_s$。故 $A\sim B$。

6　证明　先证必要性．由题设，对任意可逆阵P，有$P^{-1}AP=A$，即$AP=PA$，故$A=aI_n$。

再证充分性．由于$A=aI_n$，对任意可逆阵P，$P^{-1}AP=P^{-1}(aI_n)P=aI_n=A$．故$A$所在的相似类只有$A$本身。

7　解 1）不一定．比如本练习的第3题 AB与BA 相抵，而AB与AB既相合（当$A^{-1}=A'$时），也相似；又如

$$A=\begin{pmatrix}1&\\&1\end{pmatrix},\quad B=\begin{pmatrix}1&1\\&1\end{pmatrix}$$

显然A与B相抵，但A与B既不相合，也不相似。

2）若$A\sim B$，则存在可逆阵P，使 $B=P^{-1}AP$，而P可以写成初等阵之积，故A与B一定相抵．但不一定相合，比如

$$A=\begin{pmatrix}2&1\\-1&0\end{pmatrix},\quad B=\begin{pmatrix}1&1\\0&1\end{pmatrix},$$

A与B相似，但A与B不相合.

练　习　二

1　解

1）特征根为 $1,1,-1$；$S_{A_1}(1)=\{k_1(1,0,1)+k_2(0,1,0)\}$，$S_{A_1}(-1)=\{k(1,0,-1)\}$；属于1的特征向量为$(k_1,k_2,k_1)$，$k_1$或$k_2\neq 0$，属于$-1$的特征向量为 $(k,0,-k)$，$k\neq 0$。

2）特征根为 $1,1,-2$；$S_{A_2}(1)=\{k(3,-6,20)\}$，$S_{A_2}(-2)=\{l(0,0,1)\}$；属于1的特征向量为$(3k,-6k,20k)$，$k\neq 0$，属于$-2$的特征向量为 $(0,0,l)$，$l\neq 0$。

3）特征根为$-1,-1,-1$；$S_{A_3}(-1)=\{k(1,1,-1)\}$；属于$-1$的特征向量为$(k,k,-k)$．$k\neq 0$。

4）特征根为$2,1+\sqrt{3},1-\sqrt{3}$；$S_{A_4}(2)=\{k(2,-1,0)\}$．$S_{A_4}(1+\sqrt{3})=\{l(6+3\sqrt{3},-2-\sqrt{3},1)\}$，$S_{A_4}(1-\sqrt{3})=\{t(6-3\sqrt{3},-2+\sqrt{3},1)\}$；属于2的特征向量为$(2k,-k,0)$，

$k\neq0$，属于$1+\sqrt{3}$的特征向量为$((6+3\sqrt{3})l,\ (-2-\sqrt{3})l,l)$，$l\neq0$，属于$1-\sqrt{3}$的特征向量为$((6-3\sqrt{3})t,(-2+\sqrt{3})t,\ t)$，$t\neq0$。

5）特征根为 $2,2,2,-2$；$S_{A_5}(2)=\{k_1(1,1,0,0)+k_2(1,0,1,0)+k_3(1,0,0,1)\}$，$S_{A_5}(-2)=\{k(1,-1,-1,-1)\}$，属于2的特征向量为$(k_1+k_2+k_3,\ k_1,k_2,k_3)$，$k_1,k_2,k_3$不全为0，属于$-2$的特征向量为$(k,\ -k,\ -k,\ -k)$，$k\neq0$。

2　解　特征根全为0的阵不一定必为零阵，比如

$$A=\begin{pmatrix}0 & \cdots & 1\\ & 0 & \vdots\\ & & \ddots & \\ & & & 0\end{pmatrix}$$

的特征根全为0，但 $A\neq0$。

特征根全为1的阵不一定必为单位阵，比如

$$A=\begin{pmatrix}1 & \cdots & 1\\ & 1 & \vdots\\ & & \ddots & \\ & & & 1\end{pmatrix}$$

的特征根全为1，但A不是单位阵。

3　解　除零阵、单位阵外，还有纯量阵 aI。因为任一非零n维向量α（写为列向量的形式），有

$$(aI)\alpha=a(I\alpha)=a\alpha。$$

4　解　都不一定，比如

$$A=\begin{pmatrix}1 & \\ & 1\end{pmatrix},\quad B=\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$$

是相抵的，但A的特征多项式$(\lambda-1)^2$，B的特征多项式$\lambda^2-3\lambda-3$，即A与B的特征多项式不同；

又比如

$$A=\begin{pmatrix}1 & \\ & 1\end{pmatrix},\quad P=\begin{pmatrix}1 & 1\\ & 1\end{pmatrix},\quad B=P'AP=\begin{pmatrix}1 & \\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ & 1\end{pmatrix}=$$

$$\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}。$$

即A与B相合，但A与B的特征多项式不同。

5　解　1）可能有，比如

$$A=\begin{pmatrix}1 & \\ & 3\end{pmatrix},\quad B=\begin{pmatrix}2 & \\ & 6\end{pmatrix}$$

的特征根分别为，1,3; 2,6，故特征根不同，但容易求得A的属于1的特征向量与B的属于2的特征向量均为$(k,0)$，$k\neq 0$。

2）不能．因为若α是A的分别属于λ_1，λ_2的特征向量且$\lambda_1\neq\lambda_2$，那么由

$$A\alpha=\lambda_1\alpha,\quad A\alpha=\lambda_2\alpha,\text{有}\quad \lambda_1\alpha=\lambda_2\alpha$$

而$\alpha\neq 0$，必有$\lambda_1=\lambda_2$矛盾。

6　解　一定有．只要令$B=\lambda I$即可。

7　证明　设

$$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}.$$

由于任一n维向量都是A的特征向量，故可取

$$\alpha_1=\begin{pmatrix}1\\0\\ \vdots\\0\end{pmatrix}$$

有数λ_1，使$A\alpha_1=\lambda_1\alpha_1$，经计算得

$$a_{11}=\lambda_1,\quad a_{21}=0,\cdots,a_{n1}=0$$

同理可证，$a_{22}=\lambda_2,\quad a_{12}=0,\cdots,a_{n2}=0$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$a_{nn}=\lambda_n,\quad a_{1n}=0,\cdots,a_{n-1n}=0.$$

再取

$$\alpha=\begin{pmatrix}1\\1\\ \vdots\\1\end{pmatrix}$$

为A的特征向量，有λ，使$A\alpha=\lambda\alpha$，经计算得

$$a_{11}=a_{22}=\cdots=a_{nn}=\lambda$$

故

$$A=\begin{pmatrix}\lambda & & & \\ & \lambda & & \\ & & \ddots & \\ & & & \lambda\end{pmatrix}.$$

8 证明 $(P^{-1}AP)(P^{-1}\alpha)=P^{-1}A(PP^{-1})\alpha=P^{-1}(A\alpha)=$
$\lambda(P^{-1}\alpha)$。

9 证明 由题设知，存在α，使 $A\alpha=\lambda\alpha$。

于是

$A^m\alpha=A^{m-1}(A\alpha)=A^{m-1}(\lambda\alpha)=\lambda(A^{m-1}\alpha)=\lambda^2(A^{m-2}\alpha)=$
$\cdots=\lambda^m\alpha$，故λ^m为A^m的特征根。

10 证明

1）设λ为A的任一特征根，于是存在向量 α，使 $A\alpha=\lambda\alpha$。那么，一方面

$$A^n\alpha=0\alpha,$$

另一方面

$$A^n\alpha=A^{n-1}(A\alpha)=\lambda(A^{n-1}\alpha)=\cdots=\lambda^n\alpha$$

故 $\lambda^n\alpha=0\alpha=0$，而$\alpha\neq0$，于是$\lambda^n=0$，因此 $\lambda=0$。

3）设λ为A的特征根，存在向量α，使$A\alpha=\lambda\alpha$，那么，一方面

$$A^2\alpha=A\alpha=\lambda\alpha$$

另一方面

$$A^2\alpha=A(A\alpha)=A(\lambda\alpha)=\lambda(A\alpha)=\lambda^2\alpha$$

故 $\lambda\alpha=\lambda^2\alpha$或 $(\lambda-\lambda^2)\alpha=0$，而 $\alpha\neq0$，因此 $\lambda-\lambda^2=0$。从而 $\lambda=0$或$\lambda=1$。

2）、4）同理可证。

11 解 设A的特征根为λ，特征向量为α，于是$A\alpha=\lambda\alpha$，这里$\lambda\neq0$，否则若$\lambda=0$，可推得$\alpha=0$与α是特征向量矛盾。而由 $A\alpha=\lambda\alpha$，有$A^{-1}(A\alpha)=A^{-1}(\lambda\alpha)$，即$\alpha=\lambda(A^{-1}\alpha)$，或 $A^{-1}\alpha=\dfrac{1}{\lambda}\alpha$。于是$A^{-1}$的特征根为$A$的特征根的倒数，特征向量不变。

12　证明　令 $f(x)=a_0\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$，则 $f(A)=a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI_n$。由题设，存在数 λ_0，使 $A\alpha=\lambda_0\alpha$。于是

$$\begin{aligned}f(A)\alpha&=(a_0A^n+a_1A^{n-1}+\cdots+a_{n-1}A+aI_n)\alpha\\&=(a_0A^n)\alpha+(a_1A^{n-1})\alpha+\cdots+(a_{n-1}A)\alpha+(aI_n)\alpha\\&=a_0(A^n\alpha)+a_1(A^{n-1}\alpha)+\cdots+a_{n-1}(A\alpha)+a(I\alpha)\\&=a_0\lambda_0^n\alpha+a_1\lambda_0^{n-1}\alpha+\cdots+a_{n-1}\lambda_0\alpha+a_n\alpha\\&=(a_0\lambda_0^n+a_1\lambda_0^{n-1}+\cdots+a_{n-1}\lambda_0+a_n)\alpha\\&=f(\lambda_0)\alpha\end{aligned}$$

即α为 $f(A)$ 的属于特征根 $f(\lambda_0)$ 的特征向量。

13　证明　由题设，$A\alpha_i=\lambda\alpha_i$，$i=1,2,\cdots,m$。而

$$\begin{aligned}A(k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m)&=A(k_1\alpha_1)+A(k_2\alpha_2)+\cdots+A(k_m\alpha_m)\\&=k_1(A\alpha_1)+k_2(A\alpha_2)+\cdots+k_m(A\alpha_m)=k_1\lambda\alpha_1+k_2\lambda\alpha_2+\\&\quad\cdots+k_m\lambda\alpha_m\\&=\lambda(k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m).\end{aligned}$$

14　证明　$(\lambda I-A)'=\lambda I-A'$，有 $|(\lambda I-A)'|=|\lambda I-A'|$。而 $|(\lambda I-A)'|=|\lambda I-A|$，所以 $|\lambda I-A|=|\lambda I-A'|$。

15　证明　由

$$\lambda I_n-A=\begin{pmatrix}\lambda I_1-A_1&&&\\&\lambda I_2-A_2&&\\&&\ddots&\\&&&\lambda I_S-A_S\end{pmatrix}$$

所以

$$|\lambda I_n-A|=|\lambda I_1-A_1|\ |\lambda I_2-A_2|\cdots|\lambda I_S-A_S|.$$

16　证明　设 $S_A(\lambda_1)$，$S_A(\lambda_2)$ 且 $\lambda_1\neq\lambda_2$。任取$x\in S_A(\lambda_1)\cap S_A(\lambda_2)$，由$x\in S_A(\lambda_1)$，有$Ax=\lambda_1x$，由$x\in S_A(\lambda_2)$，有$Ax=\lambda_2x$。于是

$$\lambda_1x=\lambda_2x,\quad(\lambda_1-\lambda_2)x=0$$

若$x\neq0$，必有$\lambda_1=\lambda_2$矛盾，故 $x=0$。即

$$S_A(\lambda_1)\cap S_A(\lambda_2)=\{0\}.$$

练 习 三

1　解

1) A_1的特征向量系为$(1,0,1)$，$(0,1,0)$，$(1,0,-1)$，

$$A_1 \sim \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix},\ \text{其演化矩阵}\ P=\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix};$$

2) A_2的特征向量系为$(3,-6,20)$，$(0,0,1)$；

3) A_3的特征向量系为$(1,1,-1)$；

4) A_4的特征向量系为$(2,-1,0)$，$(6+3\sqrt{3},-2-\sqrt{3},1)$，$(6-3\sqrt{3},-2+\sqrt{3},1)$，$A_4 \sim \begin{pmatrix} 2 & & \\ & 1+\sqrt{3} & \\ & & 1-\sqrt{3} \end{pmatrix}$，演化矩阵

$$P=\begin{pmatrix} 2 & 6+3\sqrt{3} & 6-3\sqrt{3} \\ -1 & -2-\sqrt{3} & -2+\sqrt{3} \\ 0 & 1 & 1 \end{pmatrix};$$

5) A_5的特征向量系为$(1,1,0,0)$，$(1,0,1,0)$，$(1,0,0,1)$，$(1,-1,-1,-1)$，$A_5 \sim \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & -2 \end{pmatrix}$，演化矩阵 $P=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$.

2　解　结论是正确的，因为一个n阶方阵A有n个线性无关的特征向量，必有完全的特征向量系.

3　证明　必要性，设$\lambda_1,\lambda_2,\cdots,\lambda_r$是$A$的全部不同的特征根，而$S_A(\lambda_i)$的维数为$s_i$，$i=1,2,\cdots,r$. 则由题设，有

$$\Sigma s_i = n.$$

于是A有n个线性无关的特证向量.

充分性. 由题设知，每个$S_A(\lambda_i)$的维数都等于λ_i的重数，因此A是非亏损的.

4　证明　由题设知，A的每个特征空间的维数都等于其特征

根的重数，据上题，A相似于对角形；反之，不一定。如

$$A=\begin{pmatrix}2&&\\&2&\\&&\ddots&\\&&&2\end{pmatrix},$$

当然A与对角形相似，但A的特征根 2 是n重的。

5　证明　充分性．由题设

$S_A(\lambda)$的维数$=n-(\lambda I-A)$的秩$=n-(n-s)=s$。

即$S_A(\lambda)$的维数等于λ的重数，于是A与对角形相似．

必要性。A与对角形阵相似，那么对任何 s 重的特征根 λ，$S_A(\lambda)$的维数等于λ的重数等于 s 。于是

$(\lambda I-A)$的秩$=n-S_A(\lambda)$的维数$=n-s$。

6　解

$$A^k=\frac{1}{3}\begin{pmatrix}(-1)^k\cdot2+5^k & (-1)^{k+1}+5^k & (-1)^{k+1}+5^k\\(-1)^{k+1}+5^k & (-1)^k\cdot2+5^k & (-1)^{k+1}+5^k\\(-1)^{k+1}+5^k & (-1)^{k+1}+5^k & (-1)^k\cdot2+5^k\end{pmatrix}.$$

7　证明　设 $A\alpha=\lambda_1\alpha$，$A\beta=\lambda_2\beta$，$\lambda_1\neq\lambda_2$。若

$$A(k\alpha+l\beta)=\lambda(k\alpha+l\beta)$$

那么，一方面

$$\begin{aligned}A(k\alpha+l\beta)&=A(k\alpha)+A(l\beta)=k(A\alpha)+l(A\beta)\\&=k(\lambda_1\alpha)+l(\lambda_2\beta)\\&=\lambda_1(k\alpha)+\lambda_2(l\beta),\end{aligned}$$

另一方面

$$\lambda(k\alpha+l\beta)=\lambda(k\alpha)+\lambda(l\beta),$$

于是　$\lambda_1(k\alpha)+\lambda_2(l\beta)=\lambda(k\alpha)+\lambda(l\beta)$

即　$(\lambda_1-\lambda)(k\alpha)+(\lambda_2-\lambda)(l\beta)=0$

或　$(\lambda_1-\lambda)k\alpha+(\lambda_2-\lambda)l\beta=0$。

由于α，β是线性无关的，有

$$(\lambda_1-\lambda)k=0,\quad(\lambda_2-\lambda)l=0$$

又$kl\neq0$，即 $k\neq0$，$l\neq0$，那么必有

$$\lambda_1-\lambda=0,\ \lambda_2-\lambda=0\text{或}\lambda_1=\lambda=\lambda_2$$

与$\lambda_1\neq\lambda_2$矛盾，故$k\alpha+l\beta$不是A的特征向量。

8　证明　1）由题设知，A有n个不同的特征根，故A相似于对角形阵；

2）由题设知，A的特征根只有一个 $\lambda=a_{11}$，又因至少有一个$a_{i_0j_0}\neq 0\,(i_0>j_0)$，所以

$$(a_{11}I-A)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix}$$

的解空间的维数$\leqslant n-1$，说明A的特征向量系所含向量个数不超过$n-1$个，故A不与对角形阵相似。

9　解　$(\lambda-3)^2$，$(\lambda-3)^2$，$(\lambda-2)^2(\lambda-3)$。

10　证明　设$f(\lambda)$与$\varphi(\lambda)$分别为A与A'的最小多项式。由于$f(A')=(f(A))'=0$，所以$\varphi(\lambda)|f(\lambda)$。又$(\varphi(A))'=\varphi(A')=0$，有$\varphi(A)=0$，所以又有$f(\lambda)|\varphi(\lambda)$。故$f(\lambda)=\varphi(\lambda)$。

11　解　不一定。比如

$$A=\begin{pmatrix}1&\\&1\end{pmatrix},\quad B=\begin{pmatrix}1&1\\&1\end{pmatrix}$$

的特征多项式均为$(\lambda-1)^2$，而A的最小多项式为$(\lambda-1)$，B的最小多项式为$(\lambda-1)^2$。

12　证明　设诸块$A_1,A_2,\cdots,A_s$的最小多项式分别为$\varphi_1(\lambda)$，$\varphi_2(\lambda),\cdots,\varphi_s(\lambda)$。又设$f(\lambda)$为一多项式。如$f(A)=0$，则$f(A_1)=0$，$f(A_2)=0,\cdots,f(A_s)=0$。而当$f(A_i)=0$时，$\varphi_i(\lambda)\mid f(\lambda)$。故$f(\lambda)$为$\varphi_1(\lambda)$，$\varphi_2(\lambda),\cdots$，$\varphi_s(\lambda)$的公倍式。反之，如$f(\lambda)$为$\varphi_1(\lambda)$，$\varphi_2(\lambda),\cdots,\varphi_s(\lambda)$的任一公倍式，必有$f(A)=0$。因此，$A$的最小多项式为$\varphi_1(\lambda)$，$\varphi_2(\lambda),\cdots$，$\varphi_s(\lambda)$的公倍式中之次数最小者（首系数为1），即是其最小公倍式（取首系数为1）。

练 习 四

1 解 略

2 证明 必要性．由于

$$AA'=\begin{pmatrix}a_1&&\\&\ddots&\\&&a_n\end{pmatrix}\begin{pmatrix}a_1&&\\&\ddots&\\&&a_n\end{pmatrix}=\begin{pmatrix}a_1^2&&\\&\ddots&\\&&a_n^2\end{pmatrix}=I_n$$

所以，$a_i=\pm 1$，$i=1,2,\cdots,n$。

充分性，若

$$A=\begin{pmatrix}a_1&&\\&\ddots&\\&&a_n\end{pmatrix},\ \text{且}\ a_i=\pm 1,\ i=1,2,\cdots,n。$$

那么

$$A'A=\begin{pmatrix}a_1&&\\&\ddots&\\&&a_n\end{pmatrix}\begin{pmatrix}a_1&&\\&\ddots&\\&&a_n\end{pmatrix}\ \begin{pmatrix}1&&&\\&1&&\\&&\ddots&\\&&&1\end{pmatrix}$$

故

$$A'=A^{-1}。$$

3 证明 因为

$$\begin{aligned}(P^{-1}AP)(P^{-1}AP)'&=(P^{-1}AP)(P'A'P^{-1\prime})=P^{-1}APP'A'P'^{-1}\\&=P^{-1}(AA')P'^{-1}=P^{-1}P'^{-1}=I\end{aligned}$$

故$P^{-1}AP$为正交矩阵。

当A为正交矩阵，P 不是正交矩阵，$P^{-1}AP$ 仍可以是正交矩阵。例如，

$$A=\begin{pmatrix}&-1\\-1&\end{pmatrix},\quad P=\begin{pmatrix}1&\frac{1}{2}\\-1&\frac{1}{2}\end{pmatrix},\quad P^{-1}AP=\begin{pmatrix}1&\\&-1\end{pmatrix}$$

这里A为正交的，P不是正交的，但$P^{-1}AP$为正交的。

4 证明 因为

$$(P^{-1}AP)'=P'A'P^{-1\prime}=P^{-1}A(P')'=P^{-1}AP,$$

所以$P^{-1}AP$为对称阵。

5　证明　设

$$A=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}.$$

因为$A^{-1}=A'$，即

$$\begin{pmatrix}A_{11}&A_{21}&A_{31}\\A_{12}&A_{22}&A_{32}\\A_{13}&A_{23}&A_{33}\end{pmatrix}=\begin{pmatrix}a_{11}&a_{21}&a_{31}\\a_{12}&a_{22}&a_{32}\\a_{13}&a_{23}&a_{33}\end{pmatrix}$$

故得

$$a_{11}=A_{11}=a_{22}a_{33}-a_{23}a_{32}$$

$$a_{21}=A_{21}=a_{13}a_{32}-a_{12}a_{33}$$

$$a_{31}=A_{31}=a_{12}a_{23}-a_{13}a_{22}.$$

6　证明　将原方程组写成矩阵形式为

$$AX=B,\ 其中B=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

用A'左乘方程两端，得

$$X=A'B$$

即

$$x_i=(a_{1i}a_{2i}\cdots a_{ni})\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix}$$

$$=a_{1i}b_1+a_{2i}b_2+\cdots+a_{ni}b_n,\quad i=1,2,\cdots,n.$$

7　证明　设

$$A=\begin{pmatrix}a_{11}&\cdots&a_{1n}\\\cdots&\cdots&\cdots\\a_{n1}&\cdots&a_{nn}\end{pmatrix}.$$

因为A是正交矩阵，所以

$$\sum_{t=1}^{n} a_{it}a_{jt} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}.$$

由于用 -1 乘第i行各元素，那么第i行各元素平方和为

$$\sum_{t=1}^{n}(-a_{it})^2 = \sum_{t=1}^{n}a_{it}^2 = 1.$$

第i行与第 j行($i\neq j$)各对应元素相乘之和为

$$\sum_{t=1}^{n}(-a_{it})(a_{jt}) = -\sum_{t=1}^{n}a_{it}a_{jt} = 0.$$

所以，A的第i行 ($i=1,2,\cdots,n$) 乘以 -1后，仍为正交矩阵。同理可证A的第 j列 ($j=1,2,\cdots,n$)乘以 -1 后，仍为正交矩阵。

8 证明 由 1）、2）可知$A=A'$，$A'=A^{-1}$，于是 $A=A^{-1}$即$A^2=I$；由 1）、3）可知$A=A'$，$A^2=I$，即$A=A^{-1}$，于是 $A'=A^{-1}$，故A为正交矩阵；由 2）、3）可知 $A'=A^{-1}$，$A^2=I$，$A=A^{-1}$，所以$A'=A$，因此A为对称阵。

9 证明 因为

$$\begin{aligned}
&(A+S)(A-S)^{-1}[(A+S)(A-S)^{-1}]' \\
&=(A+S)(A-S)^{-1}(A-S)'^{-1}(A+S)' \\
&=(A+S)[(A'-S')(A-S)]^{-1}(A'+S') \\
&=(A+S)[(A+S)(A-S)]^{-1}(A-S) \\
&=(A+S)[(A-S)(A+S)]^{-1}(A-S) \\
&=(A+S)(A+S)^{-1}(A-S)^{-1}(A-S) \\
&=I.
\end{aligned}$$

故$(A+S)(A-S)^{-1}$为正交矩阵。

10 解

1) $e_1=(1,0,0)$，$e_2=(0,1,0)$，$e_3=(0,0,1)$。

2) $\eta_1=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}},0,0\right)$，$\eta_2=\left(\frac{1}{\sqrt{6}},-\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},0\right)$，

$$\eta_3=\left(-\frac{1}{\sqrt{12}},\ \frac{1}{\sqrt{12}},\ \frac{1}{\sqrt{12}},\ \frac{3}{\sqrt{12}}\right),\ \eta_4=\left(\frac{1}{2},\ -\frac{1}{2},\ -\frac{1}{2},\right.$$

$$\left.\frac{1}{2}\right).$$

11　解

$$P=\begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & 0 \\ -\frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

12　解　以α_1，α_2为行的矩阵为系数阵的齐次方程组

$$\begin{pmatrix} 2 & 1 & 0 & 1 & -3 \\ 1 & 1 & -1 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

的基础解系为：

$\alpha_3=(-1,2,1,0,0)$，$\alpha_4=(-1,1,0,1,0)$，$\alpha_5=(5,-7,0,0,1)$。于是$\alpha_3,\alpha_4,\alpha_5$即为所求。

13　解　基础解系为

$\alpha_1=(1,0,0,-5,-1)$，$\alpha_2=(0,1,0,-4,-1)$，$\alpha_3=(0,0,1,4,1)$。而标准正交基底为

$$e_1=\frac{1}{3\sqrt{3}}(1,0,0,-5,-1,),$$

$$e_2=\frac{1}{3\sqrt{15}}(-7,9,0,-1,-2),$$

$$e_3=\frac{1}{3\sqrt{35}}(7,6,15,1,2).$$

14　证明　因为

$$\varepsilon_i\varepsilon'_j=(t_{1i}e_1+t_{2i}e_2+\cdots+t_{ni}e_n)(t_{1j}e_1+t_{2j}e_2+\cdots+t_{nj}e_n)'$$
$$=(t_{1i}e_1+t_{2i}e_2+\cdots+t_{ni}e_n)(t_{1j}e_1'+t_{2j}e_2'+\cdots+t_{nj}e'_n)$$

$$= \sum_{h,k=1}^{n} t_{hi} t_{kj} e_h e'_k$$

$$= \sum_{s=1}^{n} t_{si} t_{sj} = \begin{cases} 1, & i=j \\ 0, & i \neq j. \end{cases}$$

所以，ε_1，ε_2，…，ε_n为标准正交基底。

15 证明 以α_1，α_2，…，α_r为行，做成一个$r \times n$矩阵T，据本节定理2，存在一个$(n-r) \times n$的矩阵T_0，使$\begin{pmatrix} T \\ T_0 \end{pmatrix}$为正交矩阵。那么$T_0$的$n-r$个行令为$\alpha_{r+1}, \cdots, \alpha_n$。故$\alpha_{r+1}, \cdots, \alpha_n$即为所求。

练 习 五

1 解

1） $e_1 = \frac{1}{3}(-2, -1, 2)$，$e_2 = \frac{1}{3}(2, -2, 1)$，$e_3 = \frac{1}{3}(1, 2, 2)$。

2） $e_1 = \left(-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right)$，$e_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0\right)$，

$e_3 = \left(-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, 0\right)$，$e_4 = \left(\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{3}{2\sqrt{3}}\right)$。

2 解

$$P = \begin{pmatrix} -\frac{1}{3} & -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & -\frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{5}{3\sqrt{5}} \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

而A为正定的，B不是正定的。

3 解

1）正交阵 $T=\frac{1}{3}\begin{pmatrix}2 & 2 & 1\\ -1 & 2 & -2\\ -2 & 1 & 2\end{pmatrix}$，平方和为 $2y_1^2-y_2^2+5y_3^2$.

2）正交阵 $T=\frac{1}{2}\begin{pmatrix}1 & -1 & -1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & 1 & 1 & 1\end{pmatrix}$，平方和为$y_1^2+7y_2^2-$

$7y_3^2-3y_4^2$。

4 证明 必要性显然成立。现在来证充分性．设 λ_1，λ_2，…，λ_n为A的特征根，也是B的特征根。那么存在正交阵T_1，T_2，使

$$T_1^{-1}AT_1=\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n\end{pmatrix}=T_2^{-1}BT_2.$$

即

$T_1^{-1}AT_1=T_2^{-1}BT_2$或$T_2T_1^{-1}AT_1T_2^{-1}=B$ 或

$(T_1T_2^{-1})^{-1}A(T_1T_2^{-1})=B$，也就是$T^{-1}AT=B$，

其中$T=T_1T_2^{-1}$。

5 证明 由于幂零阵A的特征根均为0，故存在正交阵T，使

$T^{-1}AT=0$，或 $A=TOT^{-1}=0$

6 证明 由于对称阵A的特征根均为实数，而A又是正交阵，故这些特征根的模为1．于是A的特征根均为模等于1的实数，故必为1或－1。

7 证明 由于$(\lambda I_n-A)X=0$和$(\mu I_n-A)X=0$的非零解 α和β恰分别为A的属于λ，μ的特征向量．因$\lambda\neq\mu$，故α与β正交。

8 证明 因B为正定阵，存在可逆阵C，使

$C'BC=I$。

这时，$C'AC$仍为对称阵。于是存在正交阵P，使

$$P'(C'AC)P$$

为对角形。而

$$P'(C'BC)P = P'IP = P'P = I.$$

令$T = CP$,那么

$$T'AT\text{与}T'BT$$

同时为对角形。

9 证明 由题设知，存在可逆阵P,Q,使 $A = PP'$, $B = QQ'$。于是

$$AB = (PP')(QQ') = Q'^{-1}[(P'Q)'(P'Q)]Q'.$$

其中$(P'Q)'(P'Q)$是正定的，特征根全大于0，那么AB的特征根也全大于0。

练 习 六

1 解 四阶正交阵的一切可能的标准形式为：

$$\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&1\end{pmatrix},\quad \begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&-1\end{pmatrix},\quad \begin{pmatrix}1&&&\\&1&&\\&&-1&\\&&&-1\end{pmatrix},$$

$$\begin{pmatrix}1&&&\\&-1&&\\&&-1&\\&&&-1\end{pmatrix},\quad \begin{pmatrix}-1&&&\\&-1&&\\&&-1&\\&&&-1\end{pmatrix},$$

$$\begin{pmatrix}1&&&\\&1&&\\&&\cos\theta&-\sin\theta\\&&\sin\theta&\cos\theta\end{pmatrix},\quad \begin{pmatrix}1&&&\\&-1&&\\&&\cos\theta&-\sin\theta\\&&\sin\theta&\cos\theta\end{pmatrix},$$

$$\begin{pmatrix}-1&&&\\&-1&&\\&&\cos\theta&-\sin\theta\\&&\sin\theta&\cos\theta\end{pmatrix},\quad \begin{pmatrix}\cos\theta_1&-\sin\theta_1&&\\\sin\theta_1&\cos\theta_1&&\\&&\cos\theta_2&-\sin\theta_2\\&&\sin\theta_2&\cos\theta_2\end{pmatrix}.$$

2 解 A的特征根为1，-1，$\frac{\sqrt{2}}{2}+i\frac{\sqrt{2}}{2}$，$\frac{\sqrt{2}}{2}-i\frac{\sqrt{2}}{2}$，

其标准形为

$$\begin{pmatrix} 1 & & & \\ & -1 & & \\ & & \cos\left(-\frac{\pi}{4}\right) & -\sin\left(-\frac{\pi}{4}\right) \\ & & \sin\left(-\frac{\pi}{4}\right) & \cos\left(-\frac{\pi}{4}\right) \end{pmatrix}.$$

而演化矩阵

$$T=\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

3　证明　设

$$A\alpha=\lambda_1\alpha,\quad A\beta=\lambda_2\beta,\text{ 且 }\lambda_1\neq\lambda_2$$

于是，有

$$\alpha'A'=\lambda_1\alpha',\quad \alpha'A'A\beta=(\lambda_1\alpha')\lambda_2\beta,\quad \alpha'\beta=\lambda_1\lambda_2\alpha'\beta.$$

$$(\lambda_1\lambda_2-1)\alpha'\beta=0.$$

但$(\lambda_1\lambda_2-1)\lambda_1=\lambda_1^2\lambda_2-\lambda_1=\lambda_2-\lambda_1\neq0$，故 $\lambda_1\lambda_2-1\neq0$，因之必有$\alpha'\beta=0$，从而$\alpha$与$\beta$正交。

练　习　七

1　解

(1) $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}$,　(2) $\begin{pmatrix} 0 & 0 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$,

(3) $\begin{pmatrix} 0 & -1 & & \\ 1 & 2 & & \\ & & 0 & -1 \\ & & 1 & 2 \end{pmatrix}$.

2　解

(1) $\begin{pmatrix} 1 & & \\ & 2 & \\ & & -1 \end{pmatrix}$,　(2) $\begin{pmatrix} -2 & & \\ & 1 & \\ & & 1 \end{pmatrix}$,

(3) $\begin{pmatrix} -1 & & \\ & -1 & 1 \\ & & -1 \end{pmatrix}$, (4) $\begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}$.

3 证明

1）由于A的不变因子为$d_1(\lambda)=1, d_2(\lambda)=1$, $d_3(\lambda)=\lambda^3-3\lambda^2+3\lambda-1$，所以非零次的不变因子恰为$\lambda^3-3\lambda^2+3\lambda-1$.

2）由于B的不变因子为 $d_1(\lambda)=1,\cdots,d_{m-1}(\lambda)=1$, $d_m(\lambda)=\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m$，故非零次的不变因子只有$d_m(\lambda)=\lambda^m+a_1\lambda^{m-1}+\cdots+a_{m-1}\lambda+a_m$.

4 证明

1）A的不变因子为$d_1(\lambda)=1$, $d_2(\lambda)=1$, $d_3(\lambda)=(\lambda-2)^3$，故$A$的初等因子组恰为$(\lambda-2)^3$;

2）B的不变因子为 $d_1(\lambda)=1,\cdots,d_{m-1}(\lambda)=1, d_m(\lambda)=(\lambda-\rho)^m$，故$B$的初等因子组恰为 $(\lambda-\rho)^m$.

5 解

当k为偶数时，$A^k=\begin{pmatrix} 1 & 0 & 5^k-1 \\ 0 & 5^k & 0 \\ 0 & 0 & 5^k \end{pmatrix}$,

当k为奇数时，$A^k=\begin{pmatrix} 1 & 4\cdot5^{k-1} & -1+3\cdot5^{k-1} \\ 0 & -3\cdot5^{k-1} & 4\cdot5^{k-1} \\ 0 & 4\cdot5^{k-1} & 3\cdot5^{k-1} \end{pmatrix}$.

练 习 八

1 解

1）因为$A(\lambda)$与$B(\lambda)$的不变因子均为 1，$\lambda-1$，$(\lambda-1)^2$，所以它们相抵.

2）将$A(\lambda)$经列的交换，可得

$$A_1(\lambda)=\left(\begin{array}{ccc:ccc} -1 & 0 & 0 & \lambda-a & 0 & 0 \\ 0 & -1 & 0 & 0 & \lambda-a & 0 \\ 0 & 0 & -1 & 0 & 0 & \lambda-a \\ \hdashline \lambda-a & 0 & 0 & \beta^2 & 1 & 0 \\ 0 & \lambda-a & 0 & 0 & \beta^2 & 1 \\ 0 & 0 & \lambda-a & 0 & 0 & \beta^2 \end{array}\right).$$

将 $A_1(\lambda)$ 的第1,2,3列的$(\lambda-a)$倍分别加到第4,5,6 列上，再将第1,2,3行的 $(\lambda-a)$ 倍，分别加到第4,5,6,行上，便得到 $B(\lambda)$，故 $A(\lambda)\to B(\lambda)$.

2　解　因为$A(\lambda)$的行列式因子为

$$D_1(\lambda)=1,\cdots,\ D_{n-1}(\lambda)=1,\ D_n(\lambda)=(\lambda-a)^n.$$

所以，不变因子为

$$d_1(\lambda)=1,\cdots,d_{n-1}(\lambda)=1,\ d_n(\lambda)=(\lambda-a)^n.$$

故初等因子组为$(\lambda-a)^n$.

3　证明　由$(f(x),g(\lambda))=1$，这三个 λ—矩阵各阶行列式因子均为

$$D_1(\lambda)=1,\quad D_2(\lambda)=\frac{1}{a}f(\lambda)g(\lambda),$$ 其中令$f(\lambda)g(\lambda)$的首项系数为a。故它们相抵.

4　解　$A(\lambda)$的不变因子为

$d_1(\lambda)=1,\quad d_2(\lambda)=\lambda,\quad d_3(\lambda)=(\lambda+1)(\lambda-1)\lambda^2,\quad d_4(\lambda)=(\lambda+1)^3(\lambda-1)\lambda^2,\ d_5(\lambda)=0$. 所以

$$A(\lambda)\to\begin{pmatrix}1 &&&& \\ & \lambda &&& \\ && (\lambda+1)(\lambda-1)\lambda^2 && \\ &&& (\lambda+1)^3(\lambda-1)\lambda^2 & \\ &&&& 0\end{pmatrix}.$$

5　证明

1）充分性是显然的. 下面来证必要性.

由 $A(\lambda)$ 可逆，有 $A(\lambda)=P_1(\lambda)\cdots P_S(\lambda)$，其中 $P_i(\lambda)$ 为阵，$(i=1,\cdots,s)$. 于是

$$A(\lambda)=P_1(\lambda)\cdots P_S(\lambda)I_n\text{或}P_S^{-1}(\lambda)\cdots P_1^{-1}A=I_n.$$

这说明$A(\lambda)$可经行的初等变换化为I_n。

2）证明与1）类似。

6　解

1）可逆，$\begin{pmatrix}\lambda^2 & \lambda & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}=\begin{pmatrix}1 & 0 & \lambda^2\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & \lambda & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}$.

2）可逆，并且

$$\begin{pmatrix}1 & \lambda & \lambda^3\\ 0 & 2 & \lambda\\ 0 & 0 & \frac{1}{2}\end{pmatrix}=\begin{pmatrix}1 & 0 & 2\lambda^3\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 2\lambda\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & \frac{\lambda}{2} & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & \frac{1}{2}\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

3）可逆，$\begin{pmatrix}5\lambda+1 & 25\lambda\\ \lambda & 5\lambda-1\end{pmatrix}=\begin{pmatrix}1 & 5\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \lambda & 1\end{pmatrix}\begin{pmatrix}1 & -5\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$.

4）可逆，$\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}=\begin{pmatrix}1 & 0\\ 2 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -2\end{pmatrix}\begin{pmatrix}1 & 3\\ 0 & 1\end{pmatrix}$.

7　解　n阶满秩λ—矩阵不一定可逆，比如，$A(\lambda)$是满秩的，但不可逆；$B(\lambda)$既满秩又可逆。

8　证明　由于$A(\lambda)$与$B(\lambda)$相抵，则存在可逆λ—矩阵$P(\lambda)$，$Q(\lambda)$，使

$$B(\lambda)=P(\lambda)A(\lambda)Q(\lambda).$$

于是

$$|B(\lambda)|=|P(\lambda)|\ |A(\lambda)|\ |Q(\lambda)|=a\,|A(\lambda)|\,b=ab\,|A(\lambda)|,$$

这里a，b为非零常数。

9　证明　因为

A_1的初等因子组为：$\lambda-r_1$，$\lambda-r_1$，$\lambda-r_1$；

A_2的初等因子组为：$\lambda-r_1$，$(\lambda-r_1)^2$；

A_3的初等因子组为：$(\lambda-r_1)^3$，

故A_1，A_2，A_3任两个都不相似。

10　解　因为A与B的初等因子组均为$(\lambda-1)^2$，$(\lambda-1)^2$，所以，$A\sim B$。

11　证明　因为$(\lambda I-A)'=\lambda I-A'$。所以$\lambda I-A$的行列式因子与$\lambda I-A'$的行列式因子相同，故$A\sim A'$。

12　证明　因为

$$\lambda I-A\to\begin{pmatrix} d_1(\lambda) & & \\ & \ddots & \\ & & d_n(\lambda)\end{pmatrix}$$

据上边第8题知，有

$$|\lambda I-A|=ad_1(\lambda)\cdots d_n(\lambda).$$

又因$|\lambda I-A|$是首系数为1的多项式，而$d_1(\lambda)$，…，$d_n(\lambda)$也是首系数为1的，故$a=1$。因此，有

$$|\lambda I-A|=d_1(\lambda)\cdots d_n(\lambda).$$

13　解　若当块可换为

$$\begin{pmatrix} \rho & a & & & \\ & \rho & a & & \\ & & \ddots & \ddots & \\ & & & & a \\ & & & & \rho\end{pmatrix},\quad a\neq 0.$$

因为它的初等因子组为$(\lambda-\rho)^m$。

习　题　十

1　证明

1）因为$A\sim B$，即$B=P^{-1}AP$，故A与B都可逆或都不可逆。

2）因为$A\sim B$，即$B=P^{-1}AP$，那么$B^{-1}=P^{-1}A^{-1}P$，故$A^{-1}\sim B^{-1}$。

2　证明

1）设A为幂零阵，且$A\sim B$。那么存在可逆阵P，使$B=P^{-1}AP$。由于存在自然数n，使$A^n=0$，故

$$B^n=(P^{-1}AP)^n=P^{-1}A^nP=0.$$

即 B 为幂零阵。

2）设 A 为么幂的， 且 $A\sim B$，那么存在可逆阵 P，使 $B=P^{-1}AP$。由于有自然数 n，使 $A^n=I$，故

$$B^n=(P^{-1}AP)^n=P^{-1}A^nP=P^{-1}IP=I,$$

即 B 为么幂的。

3）设 A 为幂等阵，且 $A\sim B$，那么存在可逆阵 P，使 $B=P^{-1}AP$。由于 $A^2=A$，故

$$B^2=(P^{-1}AP)^2=P^{-1}A^2P=P^{-1}AP=B,$$

即 B 也是幂等的。

3　证明　存在可逆阵 P_1，$P_2,\cdots,P_s$，使

$$B_1=P_1^{-1}A_1P_1,\quad B_2=P_2^{-1}A_2P_2,\cdots,\quad B_s=P_s^{-1}A_sP_s.$$

故

$$\begin{pmatrix}B_1&&&\\&B_2&&\\&&\ddots&\\&&&B_s\end{pmatrix}=\begin{pmatrix}P_1^{-1}A_1P_1&&&\\&P_2^{-1}A_2P_2&&\\&&\ddots&\\&&&P_s^{-1}A_sP_s\end{pmatrix}$$

$$=\begin{pmatrix}P_1^{-1}&&&\\&P_2^{-1}&&\\&&\ddots&\\&&&P_s^{-1}\end{pmatrix}\begin{pmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_s\end{pmatrix}\begin{pmatrix}P_1&&&\\&P_2&&\\&&\ddots&\\&&&P_s\end{pmatrix}.$$

$$=\begin{pmatrix}P_1&&&\\&P_2&&\\&&\ddots&\\&&&P_s\end{pmatrix}^{-1}\begin{pmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_s\end{pmatrix}\begin{pmatrix}P_1&&&\\&P_2&&\\&&\ddots&\\&&&P_s\end{pmatrix}.$$

4　证明　因为

$$|\lambda I-A|=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$$

用 $\lambda=0$ 代入，得 $|-A|=a_n$，即 $a_n=(-1)^n|A|$。

又据根与系数之关系知，$(-1)^n|A|$ 为 A 的全部特待根之积。

5　证明　因为

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

$$\xlongequal[\text{加到第一列上}]{\text{将第}2,\cdots,n\text{列都}} \begin{vmatrix} \lambda - \sum_{j=1}^{n} a_{1j} & -a_{12} & \cdots & -a_{1n} \\ \lambda - \sum_{j=1}^{n} a_{2j} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda - \sum_{j=1}^{n} a_{nj} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} 0 & -a_{12} & \cdots & -a_{1n} \\ 0 & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \cdots & \cdots & \cdots \\ 0 & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0$$

所以λ是A的特征根。又因为

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda \\ \lambda \\ \vdots \\ \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

故$(1,1,\cdots,1)$为A的属于λ的特征向量。

6　证明　设λ为A的任一特征根，存在非零的$n\times 1$矩阵α，有$A\alpha = \lambda\alpha$，或$\overline{\lambda}\overline{\alpha'} = \overline{\alpha'}\overline{A'} = -\overline{\alpha'}A$。

于是

$$\overline{\lambda}\,\overline{\alpha'}\,\alpha = -\overline{\alpha'}\,A\alpha = -\overline{\alpha'}\,(\lambda\alpha) = -\lambda\,\overline{\alpha'}\,\alpha.$$

因此，

$$(\overline{\lambda} + \lambda)\overline{\alpha'}\,\alpha = 0.$$

由于$\alpha \neq 0$，即$\overline{\alpha'}\,\alpha \neq 0$，必有$\overline{\lambda} + \lambda = 0$，故 $\lambda = 0$或λ为纯虚数。

7　证明　设λ为A的任一特征根，特征空间 $S_A(\lambda)$ 的基底为$\alpha_1, \alpha_2, \cdots, \alpha_k$。则

$$A\alpha_i = \lambda\alpha_i,\quad (i = 1, 2, \cdots, k)$$

因为$AB=BA$,所以

$$A(B\alpha_i)=(AB)\alpha_i=(BA)\alpha_i=B(A\alpha_i)=B(\lambda\alpha_i)=\lambda(B\alpha_i).$$

于是$B\alpha_i\in S_A(\lambda)$，$(i=1,2,\cdots,k)$。因而有$c_{ti}\in C$,使

$$B\alpha_i=\sum_{t=1}^{k}c_{ti}\alpha_t,\quad (i=1,2,\cdots,k).$$

现在令

$$\alpha=m_1\alpha_1+m_2\alpha_2+\cdots+m_k\alpha_k,\quad m_i\in C,\quad (i=1,2,\cdots,k).$$

即$\alpha\in S_A(\lambda)$. 如果能够适当选取m_1，$m_2,\cdots,m_k$，使$\alpha\neq 0$且$B\alpha=\mu\alpha$，$\mu\in C$，则α就是所求的A与B的公共特征向量。为此，应有

$$B\alpha=B\left(\sum_{i=1}^{k}m_i\alpha_i\right)=\sum_{i=1}^{k}m_i(B\alpha_i)=\sum_{i=1}^{k}m_i\left(\sum_{t=1}^{k}c_{ti}\alpha_t\right)$$

$$=\sum_{t=1}^{k}\left(\sum_{i=1}^{k}c_{ti}m_i\right)\alpha_t=\mu\alpha=\mu\left(\sum_{t=1}^{k}m_t\alpha_t\right)=\sum_{t=1}^{k}\mu m_t\alpha_t.$$

于是必有

$$\sum_{i=1}^{k}c_{ti}m_i=\mu m_t,\quad (t=1,2,\cdots,k).$$

即

$$(\mu I-P)\begin{pmatrix}m_1\\ \vdots\\ m_k\end{pmatrix}=\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix},$$

这里$P=(c_{ti})$是 k 阶方阵。由于在复数域C中，阵 $P=(c_{ti})$ 的特征根μ是存在的，因而阵P属于μ的特征向量

$$\begin{pmatrix}m_1\\ \vdots\\ m_k\end{pmatrix}$$

是存在的，从而A与B有公共的特征向量。

8　证明　设$d_1(\lambda)$，$d_2(\lambda),\cdots,d_n(\lambda)$为$A$的不变因子，于是据练习八的第12题，有

$$|\lambda I-A|=d_1(\lambda)d_2(\lambda)\cdots d_n(\lambda).$$

又由$d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_n(\lambda)$,所以$d_1(\lambda)$，$d_2(\lambda)$，…，$d_n(\lambda)$ 的根都是$d_n(\lambda)$的根，因而 $|\lambda I-A|$ 的根必是$d_n(\lambda)$的根。而$d_n(\lambda)$ 是A的最小多项式，故A的特征根必是A的最小多项式的根。

9　证明　设A的n个不同特征根为 $\lambda_1,\cdots,\lambda_n$。则存在可逆阵$P$，使

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

又$AB=BA$ 所以 $P^{-1}APP^{-1}BP=P^{-1}BPP^{-1}AP$，即

$$\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}(P^{-1}BP)=P^{-1}BP\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

由于$\lambda_1,\cdots,\lambda_n$互不相同，故$P^{-1}BP$为对角形。也就是$B$与对角形阵相似。

10　证明　由于

$$\begin{aligned}\varepsilon_i\varepsilon'_j&=(a_{1i}\varepsilon_1+\cdots+a_{ni}\varepsilon_n)(a_{1j}\varepsilon_1+\cdots+a_{nj}\varepsilon_n)'\\&=(a_{1i}\varepsilon_1+\cdots+a_{ni}\varepsilon_n)(a_{1j}\varepsilon_1'+\cdots+a_{nj}\varepsilon_n')\\&=\sum_{k,l=1}^{n}a_{ki}a_{lj}\varepsilon_k\varepsilon_l'\\&=\sum_{k=1}^{n}a_{ki}a_{kj}.\end{aligned}$$

而

$$\sum_{k=1}^{n}a_{ki}a_{kj}=\varepsilon_i\varepsilon_j'=\begin{cases}1, & i=j\\ 0, & i\neq j.\end{cases}$$

所以$A=(a_{ij})$为正交阵。

11　证明　因为AB^{-1}仍为正交阵，且

$$|AB^{-1}|=|A|\,|B^{-1}|=-|B|\,|B^{-1}|=-1.$$

所以，AB^{-1}必有特征根为-1，即

$$|-1\cdot I-AB^{-1}|=0,\text{或}\ |I+AB^{-1}|=0\text{或}\ |BB^{-1}+AB^{-1}|=0$$

或 $|B+A|\ |B^{-1}|=0$。因 $|B^{-1}|\neq 0$，故必 $|A+B|=0$。

12　证明

1）由于A,B都是正定的，则存在可逆阵P，使 $P'AP=I$。而$P'BP$仍为正定的，于是

$$P'(\lambda A-B)P=\lambda P'AP-P'BP=\lambda I-P'BP.$$

故

$$|\lambda A-B|=0\text{与}|\lambda I-P'BP|=0$$

有相同的根。而因$P'BP$为正定的，$|\lambda I-P'BP|=0$ 的根均为正的，从而 $|\lambda A-B|=0$的根均为正的。

2）当 $A=B$时，　$|\lambda A-B|=|\lambda I-I|\,|A|=(\lambda-1)^n|A|=0$。这时必有$(\lambda-1)^n=0$，即$\lambda=1$。故方程 $|\lambda A-B|=0$的根全为1。

反之，由1）的证明过程知，$|\lambda A-B|=0$与$|\lambda I-P'BP|=0$有相同的根，而 $|\lambda A-B|=0$的根均为1，故 $|\lambda I-P'BP|=0$的根也均为1。即$P'BP$的特征根均为1。因此，存在正交阵T，使

$$T'(P'BP)T=I\text{或}P'BP=I.$$

又知$P'AP=I$，所以有$P'AP=P'BP$，故$A=B$。

13　证明　由于λ是A的特征根，存在非零向量

$$\alpha=\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix}$$

使

$$A\alpha=\lambda\alpha.$$

把上式两边左乘α'，得

$$\alpha'A\alpha=\lambda\alpha'\alpha$$

即

$$f(\xi_1,\ \xi_2,\ \cdots,\ \xi_n)=\lambda(\xi_1^2+\xi_2^2+\cdots+\xi_n^2).$$

14　证明

1）因为S的特征根只能是零或纯虚数，所以± 1不是S的特征根，故 $|I\pm S|\neq 0$，即$I+S$，$I-S$都是可逆的。

2）因为

$$A' = (I+S)'^{-1}(I-S)' = (I-S)^{-1}(I+S).$$

而

$$(I+S)(I-S) = (I+S)I-(I+S)S = I(I+S)-S(I+S)$$
$$= (I-S)(I+S).$$

于是

$$A'A = (I-S)^{-1}(I+S)(I-S)(I+S)^{-1}$$
$$= (I-S)^{-1}(I-S)(I+S)(I+S)^{-1}$$
$$= I.$$

故$A=(I-S)(I+S)^{-1}$为正交阵。

15 证明 因A为实对称阵，所以存在正交阵T，使$T^{-1}AT$为对角形，对角线上恰为全部特征根。又因$A^2=A$，那么A的特征根只能是0或1，故

$$T^{-1}AT = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

16 证明 由题设知，A的标准形必为

$$\begin{pmatrix} 1 & & \\ & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{pmatrix},$$

其特征多项式为

$$(\lambda-1)[(\lambda-\cos\theta)^2+\sin^2\theta] = (\lambda-1)(\lambda^2-2\lambda\cos\theta+1)$$
$$= \lambda^3-(1+2\cos\theta)\lambda^2+(1+2\cos\theta)\lambda-1.$$

据Hamilton-Cayley定理，有

$$A^3-(1+2\cos\theta)A^2+(1+2\cos\theta)A-I=0.$$

取$t=1+2\cos\theta$，则$-1\leqslant t\leqslant 3$。这就证明了本题。

17 解 A,B的若当阵分别为

$$J_A=\begin{pmatrix}-1 & 1 & \\ & -1 & \\ & & 0\end{pmatrix},\quad J_B=\begin{pmatrix}1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1\\ & & 0 & 1\end{pmatrix},$$

于是与A^{100}，B^{100}相似的阵分别为

$$J_A^{100}=\begin{pmatrix}-1 & 1 & \\ & -1 & \\ & & 0\end{pmatrix}^{100}=\begin{pmatrix}1 & 100 & \\ & 1 & \\ & & 0\end{pmatrix}$$

$$J_B^{100}=\begin{pmatrix}1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1\\ & & 0 & 1\end{pmatrix}^{100}=\begin{pmatrix}1 & 100 & & \\ 0 & 1 & & \\ & & 1 & 100\\ & & 0 & 1\end{pmatrix}.$$

18　证明　必要性．A与对角形阵相似，那么A的最小多项式$\varphi(\lambda)$无重根．故取$f(\lambda)$为最小多项式$\varphi(\lambda)$即可；

充分性．设A 的最小多项式为 $\varphi(\lambda)$，由于有无重根的多项式 $f(\lambda)$，使 $f(A)=0$，则$\varphi(\lambda)\mid f(\lambda)$，从而$\varphi(\lambda)$也无重根，故 A 相似于对角形．

19　证明

1)　设A为幂零阵，即存在正整数 $m(\geqslant 2)$，有 $A^m=0$．于是可令$k(\geqslant 2)$是使$A^k=0$的最小正数．那么，$\varphi(\lambda)=\lambda^k$为$A$的最小多项式．而由$k\geqslant 2$，故$\varphi(\lambda)=\lambda^k$有重根，因此 A 不与对角形阵相似．

2)　由 1) 知，A的特征根全为 0，于是A的若当阵的对角线上的元素全为 0，即

$$J=\begin{pmatrix}0 & 1 & & & & & \\ & \ddots & \ddots & & & & \\ & & & 1 & & & \\ & & & 0 & \ddots & & \\ & & & & 0 & 1 & \\ & & & & & \ddots & \ddots \\ & & & & & & 1\\ & & & & & & 0\end{pmatrix}$$

也就是存在可逆阵P，使 $A=P^{-1}JP$．而

$$|A+I|=|P^{-1}JP+I|=|P^{-1}(J+I)P|=|P^{-1}|\,|J+I|\,|P|$$

$= |P^{-1}|\ |P|\ |J+I| = |J+I| = 1.$

20 证明 由题设知，$f(\lambda)=\lambda^m-1$以A为根. 而$f(\lambda)=\lambda^m-1$恰为m个不同的m次单位根，故没重根，据上边第18题可知A与对角形阵相似.

21 解 设A为三阶对合阵，那么A的一个化零多项式为λ^2-1，而A的最小多项式$\varphi(\lambda)$有以下三种可能：

$$\varphi(\lambda)=\lambda-1,\ \varphi(\lambda)=\lambda+1,\ \varphi(\lambda)=\lambda^2-1.$$

即A的最后一个不变因子有以下三种可能：

$$d_3(\lambda)=\lambda-1,\ d_3(\lambda)=\lambda+1,\ d_3(\lambda)=\lambda^2-1.$$

于是A的不变因子有以下四种可能：

1) $d_1(\lambda)=\lambda-1,\ d_2(\lambda)=\lambda-1,\ d_3(\lambda)=\lambda-1;$

2) $d_1(\lambda)=\lambda+1,\ d_2(\lambda)=\lambda+1,\ d_3(\lambda)=\lambda+1;$

3) $d_1(\lambda)=1,\ d_2(\lambda)=\lambda+1,\ d_3(\lambda)=\lambda^2-1;$

4) $d_1(\lambda)=1,\ d_2(\lambda)=\lambda-1,\ d_3(\lambda)=\lambda^2-1.$

故A的标准形可能有

$$\begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix},\quad \begin{pmatrix}-1&&\\&-1&\\&&-1\end{pmatrix},\quad \begin{pmatrix}-1&&\\&-1&\\&&1\end{pmatrix},$$

$$\begin{pmatrix}-1&&\\&1&\\&&1\end{pmatrix}.$$

22

$$\begin{pmatrix}a&&\\&b&\\&&c\end{pmatrix},\ \begin{pmatrix}a&&\\&b&1\\&&b\end{pmatrix},\ \begin{pmatrix}a&1&\\&a&1\\&&a\end{pmatrix}.$$

23 证明 设A的不变因子为

$$d_1(\lambda),\ d_2(\lambda),\cdots,d_n(\lambda).$$

则

$$\varphi(\lambda)=d_n(\lambda),\ D_{n-1}(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_{n-1}(\lambda).$$

由于$D_n(\lambda)=|\lambda I-A|$是关于$\lambda$的$n$次多项式，而题设$D_{n-1}(\lambda)$是$n-1$次的多项式，故

$$\varphi(\lambda)=d_n(\lambda)=\frac{D_n(\lambda)}{D_{n-1}(\lambda)}$$

是关于λ的一次式。又因为 $d_i(\lambda)\mid d_n(\lambda)$，$(i=1,2,\cdots,n-1)$。所以

$$\deg d_i(\lambda)\leqslant\deg d_n(\lambda)=1,\quad(i=1,2,\cdots,n-1).$$

又因

$$D_{n-1}(\lambda)=d_1(\lambda)d_2(\lambda)\cdots d_{n-1}(\lambda)$$

是$n-1$次的，故 $d_i(\lambda)$ 的次数不能小于1。于是 $\deg d_i(\lambda)=1$，$(i=1,2,\cdots,n-1)$。再由于 $d_i(\lambda)$ 的首系数为1且 $d_1(\lambda)\mid d_2(\lambda)\mid\cdots\mid d_{n-1}(\lambda)\mid d_n(\lambda)$。从而，有

$$d_1(\lambda)=d_2(\lambda)=\cdots=d_{n-1}(\lambda)=d_n(\lambda).$$

因此

$$D_{n-1}(\lambda)=d_1(\lambda)\cdots d_{n-1}(\lambda)=[\varphi(\lambda)]^{n-1}.$$

24 证明 必要性。由于A与对角形阵相似，所以A的任一特征根λ_i均是A的最小多项式$\varphi(\lambda)$的单根：

$$\lambda-\lambda_i\mid\varphi(\lambda),\quad(\lambda-\lambda_i)^2\nmid\varphi(\lambda).$$

于是$(\lambda-\lambda_i)^2$与 $\varphi(\lambda)$的最大公因式为$\lambda-\lambda_i$。因此存在多项式$u(\lambda)$，$v(\lambda)$，使

$$\lambda-\lambda_i=u(\lambda)(\lambda-\lambda_i)^2+v(\lambda)\varphi(\lambda).$$

故

$$A-\lambda_iI=u(A)(A-\lambda_iI)^2+v(A)\varphi(A)=u(A)(A-\lambda_iI)^2$$

于是，有

$$\begin{aligned}\text{rank}(\lambda_iI-A)&=\text{rank}(A-\lambda_iI)=\text{rank}[u(A)(A-\lambda_iI)^2]\\&=\text{rank}(u(A)(\lambda_iI-A)^2)\leqslant\text{rank}(\lambda_iI-A)^2.\end{aligned}$$

但由于

$$(\lambda_iI-A)^2=(\lambda_iI-A)(\lambda_iI-A),$$

所以，又有

$$\text{rank}(\lambda_iI-A)^2\leqslant\text{rank}(\lambda_iI-A).$$

从而，有

$$\text{rank}(\lambda_i I - A)^2 = \text{rank}(\lambda_i I - A).$$

充分性. 设λ_i是A的任一特征根. 由于$(\lambda_i I - A)^2$的秩与$(\lambda_i I - A)$的秩相同, 去证λ_i是最小多项式$\varphi(\lambda)$的单根即可.

因为λ_i是A的特征根, 必有 $\lambda - \lambda_i | \varphi(\lambda)$. 如若$\lambda_i$不是 $\varphi(\lambda)$的单根, 则一定有 $(\lambda - \lambda_i)^2 | \varphi(\lambda)$, 即$\varphi(\lambda) = (\lambda - \lambda_i)^2 g(\lambda)$. 于是 $(A - \lambda_i I) g(A) \neq 0$, 因此必有非零的$n \times 1$阵$Y$, 使

$$(A - \lambda_i I) g(A) Y \neq 0, \text{ 或 } (\lambda_i I - A) g(A) Y \neq 0.$$

但

$$(\lambda_i I - A)^2 g(A) Y = \varphi(A) Y = 0.$$

这就是说线性方程组

$$(\lambda_i I - A)^2 X = 0 \tag{1}$$

的解 $X = g(A) Y$ 不是线性方程组

$$(\lambda_i I - A) X = 0 \tag{2}$$

的解.

可另一方面, 由$\text{rank}(\lambda_i I - A) = \text{rank}(\lambda_i I - A)^2$ 知, (1)与(2)解空间的维数相同. 又(2)的解必是(1)的解, 所以(1)与(2)同解. 这就产生了矛盾, 故λ_i必是$\varphi(\lambda)$的单根.

25 证明 必要性. 因A与对角形阵相似, 由上题知, $\lambda_i I - A$的秩等于$(\lambda_i I - A)^2$的秩, 其中λ_i为A的任一特征根. 故方程组

$$(\lambda_i I - A) X = 0 \tag{1}$$

与

$$(\lambda_i I - A)^2 X = 0 \tag{2}$$

的解空间维数相等. 而(1)的解必是(2)的解, 所以(1)与(2)同解.

充分性. 由于方程组(1)与(2)同解, 故解空间的维数相同 因此, 必有

$$\text{rank}(\lambda_i I - A) = \text{rank}(\lambda_i I - A)^2.$$

据上题知, A与对角形阵相似.

26 证明 设二阶实方阵为

$$A=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}.$$

那么

$$|\lambda I-A| = \begin{vmatrix}\lambda-a_{11} & -a_{12}\\ -a_{21} & \lambda-a_{22}\end{vmatrix} = \lambda^2-(a_{11}+a_{22})\lambda+(a_{11}a_{22}-$$

$$a_{12}a_{21}).$$

由

$$|A| = a_{11}a_{22}-a_{12}a_{21}<0,$$

所以

$$\Delta=[-(a_{11}+a_{22})]^2-4(a_{11}a_{22}-a_{12}a_{21})>0.$$

从而$|\lambda I-A|$有两个不同实根λ_1，λ_2，故A相似于实对角形阵：

$$\begin{pmatrix}\lambda_1 & \\ & \lambda_2\end{pmatrix}.$$

也就是，存在实满秩阵P，使

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & \\ & \lambda_2\end{pmatrix}.$$

27　证明　设$A=(a_{ij})$，$B=(b_{ij})$，$C=(\xi_{ij})$，$|C|\neq 0$，且有关系式：

$$B=C^{-1}AC,\text{　即}CB=AC.$$

于是元素间有以下n^2个关系：

$$\xi_{i1}b_{1j}+\xi_{i2}b_{2j}+\cdots+\xi_{in}b_{nj}=a_{i1}\xi_{1j}+a_{i2}\xi_{2j}+\cdots+a_{in}\xi_{nj}$$

$$(i,\ j=1,2,\cdots,n)\qquad (1)$$

（1）可看作是关于ξ_{ij}的齐次线性方程组，由题设知（1）在复数域里有非零解，因而这个齐次线性方程组的系数阵的秩　$r<n^2$。于是有n^2-r个自由未知量，给自由未知量一组值的方法：

设r个独立未知量为$\xi_1,\cdots,\xi_r$，而自由未知量为　$\xi_{r+1},\cdots,\xi_{n^2}$，故一般解为

$$\xi_i=d_{i1}\xi_{r+1}+\cdots+d_{in^2-r}\xi_{n^2},\quad (i=1,2,\cdots,r),\qquad (2)$$

这里d_{ik}是经a_{ij}、b_{ij}加减乘得到的，所以是实数。

用（2）式代替C的元素，则有

$$|C| = \varphi(\xi_{r+1}, \xi_{r+2}, \cdots, \xi_{n^2})$$

为实数域上的n^2-r个未知量的多项式。由（1）在复数域里有非零解，可知$\varphi(\xi_{r+1}, \xi_{r+2}, \cdots, \xi_{n^2})$为非零多项式。从而一定可以在实数域中取到一组值$\xi_{r+1}^{(0)}, \cdots, \xi_{n^2}^{(0)}$，使

$$|C_0| = \varphi(\xi_{r+1}^{(0)}, \cdots, \xi_{n^2}^{(0)}) \neq 0.$$

就是说，在实数域中存在实可逆阵C_0，使

$$C_0B = AC_0, \quad 或 \quad B = C_0^{-1}AC_0.$$

第十一章　线性空间

练　习　一

1．解：由一个零向量可以构成线性空间．此外再不可能由一个向量和二个向量以及有限个向量构成线性空间

事实上，若线性空间V中有一非零向量α,则对于数域F中任意两个不同的数 k, l 都必有 $k\alpha \neq l\alpha$。因此V中若有一个非零向量就必有无限多个向量，所以除由一个零向量构成的线性空间外再没有由有限个向量构成的线性空间．

2．解：1）不是．因为两个n 次多项式相加不一定还是n 次多项式，故不是线性空间．

2）易验证该集合满足线性空间的定义的条件，故是线性空间．

3．解：不是．因为第一象限中的非零向量α，乘以 -1就属于第三象限，故不是线性空间．

4．解：1）按线性空间的定义，易验证是线性空间．

2）按线性空间定义，易验证是线性空间．

3）按线性空间定义，易验证是线性空间．

4）不是．因为两个可逆矩阵的和不一定是可逆矩阵，故不是线性空间．

5）同上可以说明不是线性空间．

5．解：不是．因为当$\alpha \neq 0$时，

$$1\cdot\alpha = 0 \neq \alpha$$

故不满足线性空间定义中算律的第8条，因此不是线性空间。

6．解：1）不是．如$1\in Z$，$\sqrt{2}\in D$，但

$$1\times\sqrt{2}=\sqrt{2}\overline{\in}Z$$

故Z不是实数域D上的线性空间．

2）同上可以说明Q不是实数域D上的线性空间．

3）按线性空间定义，易验证D是实数域D上的线性空间．

7．解：在第5题中，按其规定的集合和运算，除不满足线性空间定义中关于运算的第8条之外的一切条件都成立，因而可以说明运算的第8条是独立的，即它不能由前七条推出来．

练　习　二

1．证：1）可以验证S_1是$F^{(n)}$的子空间．

2）S_2不是$F^{(n)}$的子空间．因为$(1,0,\cdots,0)\in S_2$，$-1\in F$，但

$$-1\times(1,0,\cdots,0)=(-1,0,\cdots,0)\overline{\in}S_2$$

故S_2不是$F^{(n)}$的子空间．

2．解：不是．因为D^+虽然是D的子集，D^+与D又都是实数域D上的线性空间，但它们所规定的运算不一致，即D^+的运算不是D的运算，故D^+不是D的子空间．

3．解：不可以．因为它们不是同一数域上的线性空间，不满足子空间的定义，故不是．

4．解：不是．因为非齐次线性方程组的两个解的和不一定是非齐次线性方程组的解，故不是$F^{(n)}$的子空间．

5．证：显然$C(A)$非空，因为$I\in C(A)$．对任意的B_1，$B_2\in C(A)$，$k\in F$，则

$$(B_1+B_2)A=B_1A+B_2A=AB_1+AB_2=A(B_1+B_2)$$
$$(kB_1)A=kB_1A=kAB_1=A(kB_1)$$

故B_1+B_2，$kB_1\in C(A)$．因此，$C(A)$是$Mn(F)$的子空间．

6．证：显然$O\in L_A$，故L_A非空．对任意$B_1,B_2\in L_A$，$k\in F$，

则

$$(B_1+B_2)A=B_1A+B_2A=0+0=0$$

$$(kB_1)A=k(B_1A)=k\cdot 0=0$$

故B_1+B_2，$kB_1\in L_A$.因此L_A是$M_n(F)$的子空间。

练　习　三

1．解：在$D^{(2)}$中，令

$$W_1=\{(a,0)\mid a\in D\}$$

$$W_2=\{(0,b)\mid b\in D\}$$

显然W_1, W_2是$D^{(2)}$的子空间，可是$W_1\cup W_2$并不是$D^{(2)}$的子空间，如$(1,0)\in W_1$，$(0,1)\in W_2$但

$$(1,0)+(0,1)=(1,1)\ \overline{\in} W_1\cup W_2$$

故$W_1\cup W_2$不是$D^{(2)}$的子空间。

2．证：1）显然$S_1+(S_1\cap S_2)\supseteq S_1$，下面证明$S_1\supseteq S_1+(S_1\cap S_2)$

对任意$\alpha\in S_1+(S_1\cap S_2)$，即

$$\alpha=\alpha_1+\alpha_2$$

其中$\alpha_1\in S_1$，$\alpha_2\in S_1\cap S_2$，故$\alpha_2\in S_1$，于是

$$\alpha=\alpha_1+\alpha_2\in S_1$$

因此

$$S_1\supseteq S_1+(S_1\cap S_2)$$

所以

$$S_1=S_1+(S_1\cap S_2)$$

2）显然$S_1\supseteq S_1\cap(S_1+S_2)$，现证明 $S_1\cap(S_1+S_2)\supseteq S_1$.

对任意$\alpha\in S_1$，显然有$\alpha\in S_1+S_2$，即$\alpha\in S_1\cap(S_1+S_2)$

因此

$$S_1\cap(S_1+S_2)\supseteq S_1$$

所以

$$S_1=S_1\cap(S_1+S_2).$$

3、证：1）对任意$\alpha\in W_1\cap[(W_1\cap W_2)+W_3]$，则 $\alpha\in W_1$ 且 $\alpha\in(W_1\cap W_2)+W_3$，于是有

$$\alpha=\alpha_1+\alpha_2$$

$\alpha_1\in W_1\cap W_2$，$\alpha_2\in W_3$，故由 $\alpha\in W_1$，$\alpha_1\in W_1\cap W_2$知

$$\alpha_2=\alpha-\alpha_1\in W_1$$

即$\alpha_2\in W_1\cap W_3$，于是

$$\alpha\in(W_1\cap W_2)+(W_1\cap W_3)$$

所以

$$W_1\cap[(W_1\cap W_2)+W_3]\subseteq(W_1\cap W_2)+(W_1\cap W_3)$$

又对任意$\alpha\in(W_1\cap W_2)+(W_1\cap W_3)$，则

$$\alpha=\alpha_1+\alpha_2$$

$\alpha_1\in W_1\cap W_2$，$\alpha_2\in W_1\cap W_3$，即$\alpha_2\in W_3$，从而

$$\alpha=\alpha_1+\alpha_2\in(W_1\cap W_2)+W_3$$

即

$$\alpha\in W_1\cap[(W_1\cap W_2)+W_3]$$

于是

$$(W_1\cap W_2)+(W_1\cap W_3)\subseteq W_1\cap[(W_1\cap W_2)+W_3]$$

所以

$$(W_1\cap W_2)+(W_1\cap W_3)=W_1\cap[(W_1\cap W_2)+W_3]$$

2）对任意$\alpha\in W_1+[(W_1+W_2)\cap W_3]$，则

$$\alpha=\alpha_1+\alpha_2$$

$\alpha_1\in W_1$，$\alpha_2\in(W_1+W_2)\cap W_3$，$\alpha_2\in W_1+W_2$，$\alpha_2\in W_3$，再令

$$\alpha_2=\beta_1+\beta_2$$

$\beta_1\in W_1$，$\beta_2\in W_2$，由$\alpha_1\in W_1$，$\alpha_2\in W_3$，则有

$$\alpha=\alpha_1+\alpha_2\in W_1+W_3$$

又

$$\alpha=\alpha_1+\alpha_2=\alpha_1+(\beta_1+\beta_2)=(\alpha_1+\beta_1)+\beta_2$$

因α_1，$\beta_1\in W_1$，故 $\alpha_1+\beta_1\in W_1$，$\beta_2\in W_2$，则$\alpha\in W_1+W_2$

因此

$$\alpha\in(W_1+W_2)\cap(W_1+W_3)$$

所以

$$W_1+[(W_1+W_2)\cap W_3]\subseteq(W_1+W_2)\cap(W_1+W_3)$$

又对任意$\alpha\in(W_1+W_2)\cap(W_1+W_3)$，则$\alpha\in W_1+M_2$，$\alpha\in W_1+W_3$

故

$$\alpha=\alpha_1+\beta_1=\alpha_2+\beta_2$$

其中 $\alpha_1,\alpha_2\in W_1$，$\beta_1\in W_2$，$\beta_2\in W_3$从而

$$\beta_2=(\alpha_1-\alpha_2)+\beta_1\in W_1+W_2$$

故 $\beta_2\in(W_1+W_2)\cap W_3$，又

$$\alpha=\alpha_2+\beta_2\in W_1+[(W_1+W_2)\cap W_3]$$

因此

$$(W_1+W_2)\cap(W_1+W_3)\subseteq W_1+[(W_1+W_2)\cap W_3]$$

所以

$$(W_1+W_2)\cap(W_1+W_3)=W_1+[(W_1+W_2)\cap W_3]$$

4．证：1）对任意$\alpha\in(W_1\cap W_2)+W_3$，即

$$\alpha=\alpha_1+\alpha_2$$

其中$\alpha_1\in(W_1\cap W_2)$，$\alpha_2\in W_3$于是有 $\alpha_1\in W_1$，$\alpha_1\in W_2$故

$$\alpha=\alpha_1+\alpha_2\in W_1+W_3$$

$$\alpha=\alpha_1+\alpha_2\in W_2+W_3$$

$$\alpha\in(W_1+W_3)\cap(W_2+W_3)$$

因此

$$(W_1\cap W_2)+W_3\subseteq(W_1+W_3)\cap(W_2+W_3)$$

2）对任意$\alpha\in(W_1\cap W_3)+(W_2\cap W_3)$，即

$$\alpha=\alpha_1+\alpha_2$$

$\alpha_1\in W_1\cap W_3$，$\alpha_2\in W_2\cap W_3$，$\alpha_1\in W_3$，$\alpha_2\in W_3$，于是有

$$\alpha=\alpha_1+\alpha_2\ \in W_3$$

又 $\alpha_1\in W_1$，$\alpha_2\in W_2$，故$\alpha\in W_1+W_2$，因此

$$\alpha\in(W_1+W_2)\ \cap W_3$$

所以

$$(W_1\cap W_3)+(W_2\cap W_3)\subseteq(W_1+W_2)\cap W_3.$$

5．证：事实上，只要证明W_1不可能被W_2真包含即可。

假如 $W_1\subset W_2$，必有 $\alpha_2\in W_2$，$\alpha_2\overline{\in} W_1$但由

$$W+W_2=W+W_1$$

可有

$$\alpha+\alpha_2=\alpha'+\alpha_1 \qquad (1)$$

其中 α，$\alpha'\in W$，$\alpha_1\in W_1$由（1）可得

$$\alpha-\alpha'=\alpha_1-\alpha_2$$

又由$W_1\subset W_2$，有$\alpha_1\in W_2$因此$\alpha-\alpha'\in W$，$\alpha_1-\alpha_2\in W_2$

从而

$$\alpha-\alpha'=\alpha_1-\alpha_2\in W\cap W_2$$

又因为

$$W\cap W_1=W\cap W_2$$

因此$\alpha_1-\alpha_2\in W\cap W_1$，即$\alpha_1-\alpha_2\in W_1$又$\alpha_1\in W_1$，故$\alpha_2\in W_1$与$\alpha_2\overline{\in} W_1$矛盾，因此$W_1$不能被$W_2$真包含。即

$$W_1=W_2.$$

6．证：设

$$W=\bigcap_{\alpha} W_\alpha$$

其中W_α是既包含R又包含L的子空间，对任意$\alpha+\beta\in R+L$，$\alpha\in R$，$\beta\in L$，显然有$\alpha+\beta\in W_\alpha$，因此$\alpha+\beta\in\bigcap_\alpha W_\alpha$，即$\alpha+\beta\in W$，因此

$$R+L\subseteq W$$

又R,L皆在$R+L$中，即$R+L$也是包含R,L的子空间，因此有

$$W\subseteq R+L$$

所以

$$R+L=W$$

即$R+L$是包含R,L的最小子空间。

7．解：不成立，如平面上过原点的三条不同直线L_1,L_2,L_3分别组成三个一维子空间 W_1，W_2，W_3，显然$W_i\cap W_j=\{0\}$ $i\neq j$，

$i,j=1,2,3$。但 $W_1+W_2+W_3$不是直和。

练　习　四

1．解：当$k_1,k_2,\cdots\cdots\cdots,k_r$中有为 0 的数时，显然$k_1\alpha_1,k_2\alpha_2,\cdots,k_r\alpha_r$中有零向量，则一定线性相关。

反之，若$k_1,k_2,\cdots,k_r$全不为 0 时，由等式

$$l_1(k_1\alpha_1)+l_2(k_2\alpha_2)+\cdots+l_r(k_r\alpha_r)=0$$

$$(l_1k_1)\alpha_1+(l_2k_2)\alpha_2+\cdots+(l_rk_r)\alpha_r=0$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_r$线性无关，故必有

$$l_ik_i=0\quad i=1,2,\cdots,r$$

因为$k_i\neq 0$，故l_i必为 0 ，因此$k_1\alpha_1,k_2\alpha_2,\cdots,k_r\alpha_r$线性无关。

2．证：由三角公式可知

$$\cos 2t=2\cos^2 t-1$$

即

$$1+\cos 2t-2\cos^2 t=0$$

故1，$\cos^2 t$，$\cos 2t$ 线性相关。

3．由等式

$$k_1\cdot 1+k_2(1+x)+\cdots+k_n(1+x+\cdots+x^{n-1})=0$$

有

$$(k_1+k_2+\cdots+k_n)+(k_2+\cdots+k_n)x+\cdots+k_nx^{n-1}=0$$

因 $1,x,\cdots,x^{n-1}$ 是线性无关的，故有

$$\begin{aligned}k_1+k_2+\cdots+k_n&=0\\ k_2+\cdots+k_n&=0 \qquad (1)\\ \cdots\cdots\cdots\cdots&\cdots\cdots\\ k_n&=0\end{aligned}$$

显然方程组（1）有且仅有零解，故 $1,1+x,\cdots,1+x+\cdots+x^{n-1}$ 是线性无关的。

4．证：（反证法）设 $f_1(x),f_2(x)$，$f_3(x)$ 是线性相关的，则

有不全为 0 的数k_1, k_2, k_3使等式

$$k_1 f_1(x) + k_2 f_2(x) + k_3 f_3(x) = 0$$

成立，不妨设$k_1 \neq 0$，于是有

$$f_1(x) = -\frac{k_2}{k_1} f_2(x) - \frac{k_3}{k_1} f_3(x) \qquad (1)$$

又因 $f_2(x)$，$f_3(x)$不互质，则必有非零次多项式$d(x)$整除 $f_2(x)$，$f_3(x)$，由（1）式立即可得

$$d(x) \mid f_1(x)$$

即 $d(x)$ 是 $f_1(x)$，$f_2(x)$，$f_3(x)$ 的公因式，这与 $f_1(x)$，$f_2(x)$，$f_3(x)$互质相矛盾，因此，$f_1(x)$，$f_2(x)$，$f_3(x)$是线性无关的。

5．证：考查等式

$$l_1\beta_1 + l_2\beta_2 + \cdots + l_r\beta_r = 0$$

即

$$l_1\alpha_1 + l_2\alpha_2 + \cdots + l_{r-1}\alpha_{r-1} + (l_1k_1 + l_2k_2 + \cdots + l_{r-1}k_{r-1} + l_r)\alpha_r = 0$$

由 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关有$l_1, l_2, \cdots, l_{r-1}, l_1k_1 + l_2k_2 + \cdots + l_{r-1}k_{r-1} + l_r$必为 0，从而有$l_1, l_2, \cdots, l_{r-1}, l_r$全等于 0，故$\beta_1, \beta_2, \cdots, \beta_r$是线性无关的。

6．证：事实上，若证$\alpha_1, \alpha_2, \cdots, \alpha_r$与 $\alpha_1, \alpha_2, \cdots, \alpha_{r-1}, \beta$等价，只须证$\alpha_r$可被$\alpha_1, \alpha_2, \cdots, \alpha_{r-1}, \beta$线性表出即可。

因β 可被 $\alpha_1, \alpha_2, \cdots, \alpha_r$线性表出，于是有

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_{r-1}\alpha_{r-1} + k_r\alpha_r$$

显然$k_r \neq 0$，若 $k_r = 0$，则β 可被 $\alpha_1, \alpha_2, \cdots, \alpha_{r-1}$ 线性表出，这与题设矛盾．于是

$$\alpha_r = \frac{1}{k_r}\beta - \frac{k_1}{k_r}\alpha_1 - \frac{k_2}{k_r}\alpha_2 - \cdots - \frac{k_{r-1}}{k_r}\alpha_{r-1}$$

即α_r 被$\alpha_1, \alpha_2, \cdots, \alpha_{r-1}, \beta$线性表出，因此可得 $\alpha_1, \alpha_2, \cdots, \alpha_r$与 $\alpha_1, \alpha_2, \cdots \alpha_{r-1}, \beta$，是等价的。

7．证：考查向量组

$$\alpha_1, \alpha_2, \cdots, \alpha_r, \beta \qquad (1)$$

$$\alpha_1, \alpha_2, \cdots, \alpha_r, \gamma \qquad (2)$$

若（1）与（2）中至少有一组是线性相关时，由$\alpha_1, \alpha_2, \cdots, \alpha_r$线性无关则立即可有 β, γ中至少有一个被 $\alpha_1, \alpha_2, \cdots, \alpha_r$线性表出。

若（1）与（2）都线性无关，由 $\alpha_1, \alpha_2, \cdots, \alpha_r, \beta, \gamma$ 线性相关可知β 可被（2）线性表出，γ 可被（1）线性表出，故（1）与（2）等价。

练 习 五

1．解：1）考查等式

$$k_1\alpha_1 + k_2\alpha_2 + k_3\alpha_3 + k_4\alpha_4 = 0$$

$$k_1(x^3+1) + k_2(x+1) + k_3(x^2+x) + k_4(x^3+x^2+2x+2) = 0$$

$$(k_1+k_2+2k_4) + (k_2+k_3+2k_4)x + (k_3+k_4)x^2 + (k_1+k_4)x^3 = 0$$

由$1, x, x^2, x^3$线性无关有

$$\begin{cases} k_1 + k_2 + 2k_4 = 0 \\ k_2 + k_3 + 2k_4 = 0 \\ k_3 + k_4 = 0 \\ k_1 + k_4 = 0 \end{cases} \qquad (1)$$

由（1）有非零解，故$\alpha_1, \alpha_2, \alpha_3, \alpha_4$线性相关。因此不是$F_4[x]$ 的基底。

2）考查等式 $k_1\beta_1 + k_2\beta_2 + k_3\beta_3 + k_4\beta_4 = 0$

$$k_1(1-x^2) + k_2(x-1) + k_3(x^2+2x-2) + k_4x^3 = 0$$

$$(k_1-k_2-2k_3) + (k_2+2k_3)x + (k_3-k_1)x^2 + k_4x^3 = 0$$

由于$1, x, x^2, x^3$是线性无关，则有

$$\begin{cases} k_1 - k_2 - 2k_3 = 0 \\ k_2 + 2k_3 = 0 \\ k_3 - k_1 = 0 \\ k_4 = 0 \end{cases} \qquad (2)$$

因（2）只有零解，故$\beta_1,\beta_2,\beta_3,\beta_4$是线性无关的，因此是$F_4[x]$的基底。

2．解：显然

$$D^{(4)}=L(\alpha_1,\alpha_2,\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4)$$

其中$\varepsilon_1=(1,0,0,0),\varepsilon_2=(0,1,0,0),\varepsilon_3=(0,0,1,0),\varepsilon_4=(0,0,0,1)$。向量组

$$\alpha_1,\alpha_2,\ \varepsilon_1,\varepsilon_2,\ \varepsilon_3,\varepsilon_4, \qquad (1)$$

是$D^{(4)}$的生成组．在（1）中依次由左往右挑选4个不能由前面的向量线性表出的向量，即是$D^{(4)}$的基底，易验证α_1，α_2，$\varepsilon_1,\varepsilon_2$为所求，即是由$\alpha_1,\alpha_2$扩充成为$D^{(4)}$的基底。

3．解：令F_{ij}表示i行j列位置的元素和j行i列位置上的元素为1，而其余位置上的元素全为0的n阶方阵，易验证F_{ij} $(i,j=1,2,\cdots,n)$是S的基底，故S的维数是

$$n+(n-1)+\cdots+2+1=\frac{n(n+1)}{2}$$

4、证：因为$1,i$是C的基底，因此C在实数域上构成的线性空间的维数是2。

当C看做自身上线性空间，则1是它的基底，故维数是1。

5．证：因为V中的每个向量都可被$\alpha_1,\alpha_2,\cdots\alpha_n$唯一地的线性表出，则$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$的生成组，而0向量表法唯一，故$\alpha_1$，$\alpha_2,\cdots$，$\alpha_n$是线性无关的，从而是$V$的基底，所以

$$\dim V=n.$$

6．证：若证$\alpha_1,\alpha_2,\cdots$，α_n是V的基底，只须证$\alpha_1,\alpha_2,\cdots,\alpha_n$线性无关即可

由β可被$\alpha_1,\alpha_2,\cdots,\alpha_n$唯一地线性表出，则有

$$\beta=l_1\alpha_1+l_2\alpha_2+\cdots+l_n\alpha_n \qquad (1)$$

假如$\alpha_1,\alpha_2,\cdots$，α_n线性相关，则有不全为0的数$k_1,k_2,\cdots$，k_n使得下式成立

$$0=k_1\alpha_1+{}_2\alpha_2+\cdots+k_n\alpha_n \qquad (2)$$

将（1），（2）两式相加得

$$\beta=(k_1+l_1)\alpha_1+(k_2+l_2)\alpha_2+\cdots+(k_n+l_n)\alpha_n \qquad (3)$$

由于$k_1,k_2,\cdots,k_n$不全为0，则(3)与（1）式中的系数不全相同，这与β的表法唯一矛盾．故$\alpha_1,\alpha_2,\cdots,\alpha_n$线性无关，因此$\alpha_1,\alpha_2,\cdots\alpha_n$是$V$的基底。

7．证：在W中任取$\alpha=(a_1,a_2,\cdots,a_n)\neq 0$，而令$\beta=(b_1,b_2,\cdots,b_n)$是$W$中任意向量．

当$\beta=0$时，显然β可被α线性表出．

当$\beta\neq 0$时，假如β不能被α线性表出，

即对任何k都有$\beta\neq k\alpha$，即$\beta-k\alpha\neq 0$。

因为$\alpha\neq 0$，所以至少有一个分量不为0，不妨设$a_1\neq 0$，于是取

$$k=\frac{b_1}{a_1}\text{有}$$

$$\beta-\frac{b_1}{a_1}\alpha=(0,*,\cdots,*)\neq 0$$

因β，$\alpha\in W$，故$\beta-\frac{b_1}{a_1}\alpha\in W$，与题设矛盾，故有$\beta=k\alpha$，又$\alpha\neq 0$，则$\alpha$是$W$的基，则$\dim W=1$。

练　习　六

1．解：同一向量在不同基底上的坐标有时是可以相同的．如零向量在任何基底上的坐标都相同．再如$\varepsilon_1=(1,0,0)$，$\varepsilon_2=(0,1,0)$，$\varepsilon_3=(0,0,1)$是$F^{(3)}$的一个基底，从而ε_1，$2\varepsilon_2$，$3\varepsilon_3$；$\varepsilon_1,\varepsilon_3$，$\varepsilon_2$也是$F^{(3)}$的基底．于是向量$\alpha=(1,0,0)=\varepsilon_1$在这三个基底上的坐标都是1，0，0．

当基底选定之后，每一向量的坐标是唯一确定的，所以不同向量在同一基底上的坐标一定不同。

2．解：易验证向量组

$$1, x, \cdots, x^{n-1}, (x+a)^n$$

与

$$1,\ x+a, \cdots, (x+a)^{n-1}, (x+a)^n$$

是$F_{n+1}[x]$的两组基底，显然$g(x)=(x+a)^n$在这两组基底上的坐标皆为0，0，……，1．故相同．则这二基底为所求．

3．证：欲证x^2+x，x^2-x，$x+1$是$F_3[x]$的基底只须证明它们是线性无关即可．

考查等式

$$k_1(x^2+x)+k_2(x^2-x)+k_3(x+1)=0$$

$$(k_1+k_2)x^2+(k_1-k_2+k_3)x+k_3=0$$

由于 x^2，x，1是线性无关的，于是有

$$\begin{cases} k_1+k_2=0 \\ k_1-k_2+k_3=0 \\ k_3=0 \end{cases} \qquad (1)$$

解（1）知，只有零解，故x^2+x，x^2-x，$x+1$是线性无关，故是$F_3[x]$的基底．

下面求$2x^2+7x+3$的坐标，即

$$k_1(x^2+x)+k_2(x^2-x)+k_3(x+1)=2x^2+7x+3$$

于是有

$$\begin{cases} k_1+k_2=2 \\ k_1-k_2+k_3=7 \\ k_3=3 \end{cases}$$

解得 $k_1=3$，$k_2=-1$，$k_3=3$．故向量 $2x^2+7x+3$在基底 x^2+x，x^2-x，x^2+1上的坐标是3、-1、3．

4．解：

$$\alpha_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\ \alpha_2=\begin{pmatrix}0&1\\0&0\end{pmatrix},\ \alpha_3=\begin{pmatrix}0&0\\1&0\end{pmatrix},\ \alpha_4=\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

$$\beta_1=\begin{pmatrix}1&0\\0&0\end{pmatrix},\ \beta_2=\begin{pmatrix}1&1\\0&0\end{pmatrix},\ \beta_3=\begin{pmatrix}1&1\\1&0\end{pmatrix},\ \beta_4=\begin{pmatrix}1&1\\1&1\end{pmatrix}$$

显然$\{\alpha_i\}$与$\{\beta_i\}$是$M_2(F)$的两组基底，而且A在$\{\alpha_i\}$上的坐标为-1，$3, 0, 2$. A在$\{\beta_i\}$上的坐标为$-4, 3, -2, 2$

5．解：由于

$$\alpha_2 = 0\cdot\alpha_1 + 1\cdot\alpha_2 + 0\cdot\alpha_3 + \cdots + 0\cdot\alpha_n$$

$$\alpha_3 = 0\cdot\alpha_1 + 0\cdot\alpha_2 + 1\cdot\alpha_3 + \cdots + 0\cdot\alpha_n$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\alpha_n = 0\cdot\alpha_1 + 0\cdot\alpha_2 + 0\cdot\alpha_3 + \cdots + 1\cdot\alpha_n$$

$$\alpha_1 = 1\cdot\alpha_1 + 0\cdot\alpha_2 + 0\cdot\alpha_3 + \cdots + 0\cdot\alpha_n$$

因此由基底$\alpha_1, \alpha_2, \cdots\alpha_n$到基底$\alpha_2, \alpha_3\cdots\alpha_n, \alpha_1$的过渡阵为

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

6．解：1）易得由$\{\varepsilon_i\}$到$\{\eta_i\}$的过渡阵为

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

2）设所求的基底为α_1，α_2，α_3，于是有

$$(\alpha_1, \alpha_2,\ \alpha_3) = (\eta_1\eta_2\eta_3)A$$

即

$$\alpha_1 = \eta_1 = (1, 0, 0)$$

$$\alpha_2 = \eta_2 - \eta_1 = (1, 1, 0) - (1, 0, 0) = (0, 1, 0)$$

$$\alpha_3 = \eta_3 - \eta_2 = (1, 1, 1) - (1, 1, 0) = (0, 0, 1)$$

故所得的新基底为

$$\alpha_1 = (1, 0, 0),\ \ \alpha_2 = (0, 1, 0),\ \ \alpha_3 = (0, 0, 1)$$

7．解：1）易求得由基底（1）到基底（2）的过渡阵为

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

2）易求得由基底（2）到基底（1）的过渡阵为

$$A^{-1}=\begin{pmatrix}1&-1&0&0\\0&1&-1&0\\0&0&1&-1\\0&0&0&1\end{pmatrix}$$

由$g(x)$在基底（1）的坐标为1,0,－2,5可求得$g(x)$在基底（2）的坐标为

$$A^{-1}\begin{pmatrix}1\\0\\-2\\5\end{pmatrix}=\begin{pmatrix}1\\2\\-7\\5\end{pmatrix}$$

故$g(x)$在（2）上的坐标为1,2,－7,5。再由$f(x)$在（2）上的坐标为7,0,8,－2可得$f(x)+g(x)$在（2）上的坐标为8,2,1,3。

由于$f(x)$在基底（2）上的坐标为7,0,8,－2可求得$f(x)$在基底（1）上的坐标为

$$A\begin{pmatrix}7\\0\\8\\-2\end{pmatrix}=\begin{pmatrix}13\\6\\6\\-2\end{pmatrix}$$

故$f(x)$在（1）上的坐标为13,6,6,－2,而$g(x)$在（1）上的坐标为1,0,－2,5,可得$f(x)+g(x)$在基底（1）上的坐标为14,6,4,3。

练 习 七

1．证：因为容易验证向量组

$$\begin{pmatrix}1&0\\0&0\end{pmatrix},\ \begin{pmatrix}0&1\\0&0\end{pmatrix},\ \begin{pmatrix}0&0\\1&0\end{pmatrix},\ \begin{pmatrix}0&0\\0&1\end{pmatrix}$$

$$\begin{pmatrix}i&0\\0&0\end{pmatrix},\ \begin{pmatrix}0&i\\0&0\end{pmatrix},\ \begin{pmatrix}0&0\\i&0\end{pmatrix},\ \begin{pmatrix}0&0\\0&i\end{pmatrix}$$

是$M_2(C)$的基底，故$M_2(C)$是8维线性空间，而$D^{(4)}$也是8维线性空间，即$M_2(C)$与$D^{(4)}$它们的维数相同，因此它们同构。

2．证：显然D是1维线性空间，下面只须证明D^+也是一维线性空间即可

事实上设$a\neq 1\in D^+$，显然线性无关，对任意$b\in D^+$，都有$\log_a b\in D$，使得

$$\log_a b\circ a=a^{\log_a b}=b$$

即a是D^+的生成组，故a是D^+的基底，因此D^+是1维的，所以D与D^+同构。

3．证：令$\bar{F}[x]$是数域F上常数项等于0的多项式全体所构成的线性空间，显然$\bar{F}[x]$是$F[x]$的真子空间，下面可以证明

$$F[x]\simeq\bar{F}[x]$$

事实上，令

$$\varphi: f(x)\longmapsto xf(x)$$

显然当$f(x)\neq g(x)$时，$xf(x)\neq xg(x)$，故φ是单射，而且对任意$h(x)\in\bar{F}[x]$都

$$h(x)=xh_1(x)$$

其中$h_1(x)\in F[x]$，故φ是满射．即为单满射．对任意$f(x)$，$g(x)\in F[x]$，都有

$$\begin{aligned}\varphi(f(x)+g(x))&=x(f(x)+g(x))\\&=xf(x)+xg(x)\\&=\varphi(f(x))+\varphi(g(x))\end{aligned}$$

$$\begin{aligned}\varphi(k\cdot f(x))&=x(kf(x))\\&=k(xf(x))\\&=k\varphi(f(x))\end{aligned}$$

因此

$$F[x]\simeq\bar{F}[x].$$

习 题 十一

1．证：由规定易验证加法满足交换律，结合律．

而 $(0,0)\in M$，它是M的零元．

对任意$(a,b)\in M$，则$(-a,\ a^2-b)\in M$，它是(a,b)的负元．

对任意$k, l \in D$，$(a, b) \in M$

$$(k+l)\cdot(a, b) = ((k+l)a, \frac{1}{2}(k+l)(k+l-1)a^2 + (k+l)b)$$

$$k\circ(a, b) \oplus l\circ(a, b) = (ka, \frac{1}{2}k(k-1)a^2 + kb) \oplus$$

$$(la, \frac{1}{2}l(l-1)a^2 + lb)$$

$$= (ka + la, \frac{1}{2}k(k-1)a^2 + kb$$

$$+ \frac{1}{2}l(l-1)a^2 + lb + ka\cdot la)$$

$$= ((k+l)a, \frac{1}{2}(k+l)$$

$$(k+l-1)a^2 + (k+l)b)$$

故

$$(k+l)\circ(a, b) = k\circ(a, b) \oplus l\circ(a, b)$$

对任意$k \in D$，$(a_1, b_1), (a_2, b_2) \in M$

$$k\circ((a_1, b_1) \oplus (a_2, b_2)) = k\circ(a_1 + a_2, b_1 + b_2 + a_1 a_2)$$

$$= (k(a_1 + a_2), \frac{1}{2}k(k-1)(a_1 + a_2)^2$$

$$+ k(b_1 + b_2 + a_1 a_2))$$

$$k\circ(a_1, b_1) \oplus k\circ(a_2, b_2) = (ka_1, kb_1 + \frac{1}{2}k(k-1)a_1^2) \oplus$$

$$(ka_2, kb_2 + \frac{1}{2}k(k-1)a_2^2)$$

$$= (ka_1 + ka_2, kb_1 + \frac{1}{2}k(k-1)a_1^2 + kb_2 + \frac{1}{2}k(k-1)a_2^2$$

$$+ ka_1 ka_2)$$

$$= (k(a_1 + a_1), k(b_1 + b_2 + a_1 a_2) + \frac{1}{2}k(k-1)(a_1^2 + a_2^2$$

$$+ 2a_1 a_2))$$

$$= (k(a_1 + a_2), k(b_1 + b_2 + a_1 a_2) + \frac{1}{2}k(k-1)(a_1 + a_2)^2)$$

故

$$k\circ((a_1,b_1)\oplus(a_2,b_2))=k\circ(a_1,b_1)\oplus k\circ(a_2,b_2)$$

对任意 $k,l\in D,(a,b)\in M$,

$$k\circ(l\circ(a,b))=k\circ(la,\ \frac{1}{2}l(l-1)a^2+lb)$$

$$=(kla,k(lb+\frac{1}{2}l(l-1)a^2)+\frac{1}{2}k(k-1)(la)^2)$$

$$=(kla,\ \frac{1}{2}kl(kl-1)a^2+klb)$$

$$=(kl)\circ(a,b)$$

$$1\circ(a,b)=(1\cdot a,1\cdot b+\frac{1}{2}1\cdot(1-1)a^2)$$

$$=(a,b)$$

因此M是线性空间。

2．解：不是。当$0\neq\alpha\in\pi$

$$(1+1)\circ\alpha=2\circ\alpha=\alpha$$

$$1\circ\alpha+1\circ\alpha=\alpha+\alpha=2\alpha$$

显然

$$\alpha\neq2\alpha$$

故π 不是线性空间。

3．证：1）因为$\alpha\in W_2$，故$k\alpha\in W_2$，则$\beta+k\alpha$就一定不属于W_2。

事实上，令$\beta+k\alpha=\gamma\in W_2$，再由$k\alpha\in W_2$，则

$$\beta=\gamma-k\alpha\in W_2$$

这与$\beta\overline{\in}W_2$矛盾，故$\beta+k\alpha\overline{\in}W_2$。

2）当 $\beta\in W_1$时，因$\alpha\overline{\in}W_1$，当$k\neq0$时$k\alpha\overline{\in}W_1$，则 $\beta+k\alpha\overline{\in}W_1$，因此只有$k=0$时$\beta+k\alpha\ \overline{\in}W_1$。

当 $\beta\overline{\in}W_1$时，又因$k\alpha\ \overline{\in}W_1$，则有可能存在$k$ 使得$\beta+k\alpha\in W_1$，下面证明若存在这样的k 也只能是唯一的。

假如有$k_1\neq k_2$，使得$\beta+k_1\alpha\in W_1$，$\beta+k_2\alpha\in W_1$，

于是

$$(\beta+k_1\alpha)-(\beta+k_2\alpha)\in W_1$$

即

$$(k_1-k_2)\alpha\in W_1$$

由于$k_1-k_2\neq 0$ 从而得 $\alpha\in W_1$，这与题设相矛盾。所以至多有一个$k\in D$，使得$\beta+k\alpha\in W_1$。

4．证：因W_1, W_2 是V的真子空间，故有V中向量 $\alpha'\overline{\in} W_1$，$\beta'\overline{\in} W_2$，如果

i）当$\alpha'\overline{\in} W_2$时，则 $\alpha=\alpha'$即为所求。

ii）当 $\alpha'\in W_2$，但 $\beta'\overline{\in} W_1$时，$\alpha=\beta'$即为所求。

iii）当$\alpha'\overline{\in} W_1$但 $\alpha'\in W_2$，$\beta'\overline{\in} W_2$但$\beta'\in W_1$时，

那么$\alpha=\alpha'+\beta'$即为所求。

事实上，若$\alpha\in W_1$，由$\beta'\in W_1$那么$\alpha'=\alpha-\beta'\in W_1$，这与$\alpha'\overline{\in} W_1$矛盾，故 $\alpha\overline{\in} W_1$。同理可证 $\alpha\overline{\in} W_2$，因此α 同时不属于W_1, W_2。即为所求。

5．证：（归纳法）当$r=2$时由上题知成立．假设$r=k-1$ 时也成立，即存在$\alpha\in V$使得α 同时不属于W_i，$(i=1,2,\cdots,k-1)$下面证明$r=k$时也成立．显然当此 $\alpha\overline{\in} W_k$则α 即为所求。若$\alpha\in W_k$，由于W_k是V的真子空间，故有$\beta\overline{\in} W_k$，再由 3 题2)知对任意 W_i $(i=1,2,\cdots,k-1)$都至多有一个k_i使得

$$\beta+ki\alpha\in W_i \quad i=1,2,\cdots,k-1,$$

我们取异于k_i的数 k，则必有

$$\beta+k\alpha\overline{\in} W_i, i=1,2,\cdots,k-1$$

又由于$\alpha\in W_k$，$\beta\overline{\in} W_k$故有 $\beta+k\alpha\ \overline{\in} W_k$，故$\beta+k\alpha$同时不属于$W_i$$(i=1,2,\cdots,k)$。

6．令$W=L(1,\sqrt{2},\sqrt{5})$，显然 1，$\sqrt{2},\sqrt{5}$ 是W的生成组．考查等式

$$a\cdot 1+b\cdot\sqrt{2}+c\cdot\sqrt{5}=0 \quad a,b,c\in Q$$

则有

$$(a+b\sqrt{2})^2=(-c\sqrt{5})^2$$
$$a^2+2ab\sqrt{2}+2b^2=5c^2$$
$$2ab\sqrt{2}=5c^2-a^2-2b^2\in Q$$
所以$2ab=0$，从而$a=0$或$b=0$。

若$a=0$，则有
$$b\sqrt{2}+c\sqrt{5}=0$$
$$\frac{c}{2}\sqrt{10}=-b\in Q$$
所以$c=0$，从而$b=0$

若$b=0$，则有
$$a+c\sqrt{5}=0$$
$$c\sqrt{5}=-a\in Q$$
所以$c=0$，从而$a=0$，即1，$\sqrt{2}$，$\sqrt{5}$ 线性无关，是W的基底，所以$L(1,\sqrt{2},\sqrt{5})$的维数是3。

7．由题设（1）有 $W_1\cap W_2\subset W_1+W_2$ 由此可证得，$W_1\neq W_2$。

事实上，若 $W_1=W_2$，则 $W_1\cap W_2=W_1$，$W_1+W_2=W_1$ 这与$W_1\cap W_2\subset W_1+W_2$矛盾，故$W_1\neq W_2$。

显然 $W_1\cap W_2\subseteq W_1$，$W_1\cap W_2\subseteq W_2$。同时进一步可以证明这里必有一个真包含。

事实上若不存在真包含，则有
$$W_1\cap W_2=W_1,\quad W_1\cap W_2=W_2$$
进而$W_1=W_2$，这与$W_1\neq W_2$矛盾，故必有一个真包含。不妨设$W_1\cap W_2\subset W_1$，于是有
$$\dim W_1>\dim(W_1\cap W_2)$$
又因为$W_1\subseteq W_1+W_2$，再由题设（1）则必有
$$W_1=W_1+W_2$$
又因为$W_1\neq W_2$，则 $W_2\neq W_1+W_2$，即 $W_2\subset W_1+W_2$于是
$$\dim W_2<\dim(W_1+W_2)$$

再由$W_1 \cap W_2 \subseteq W_2$和依据题设（1）可有

$$\dim W_2 = \dim(W_1 \cap W_2)$$

故

$$W_2 = W_1 \cap W_2$$

从而得证。

8．证：因为

$$\deg f^{(i)}(x) = n - i,\quad i = 0, 1, \cdots, n$$

故

$$f(x), f'(x), \cdots, f^{(n)}(x)$$

是线性无关的，又因为$F_{n+1}[x]$是$n+1$维的，由此可得$f(x)$，$f'(x)$，$\cdots$，$f^{(n)}(x)$是 $F_{n+1}[x]$的基底。

如果

$$f(x) = (x-a)^n$$

那么

$$f^{(i)}(x) = \frac{n!}{(n-i)!}(x-a)^{n-i} \qquad i = 0, 1, \cdots, n$$

对任意$g(x) \in F_{n+1}[x]$，令

$$g(x) = k_1 f(x) + k_2 f'(x) + \cdots + k_{n+1} f^{(n)}(x)$$

$$= k_1(x-a)^n + k_2 n(x-a)^{n-1} + \cdots + ki\frac{n!}{(n-i)!}(x-a)^{n-i}$$

$$+ \cdots + k_{n+1} n!$$

令$x = a$有

$$g(a) = k_{n+1} n!$$

故

$$k_{n+1} = \frac{g(a)}{n!}$$

仿此可有

$$g'(a) = n!\, k_n$$

$$k_n = \frac{g'(a)}{n!}$$

$$\vdots$$

$$k_2 = \frac{g^{(n-1)}(a)}{n!}$$

$$k_1 = \frac{g^{(n)}(a)}{n!}$$

于是$g(x)$在基底$f(x), f'(x), \cdots, f^{(n)}(x)$上的坐标为

$$\frac{g^{(n)}(a)}{n!}, \frac{g^{(n-1)}(a)}{n!}, \cdots, \frac{g'(a)}{n!}, \frac{g(a)}{n!}.$$

9．证：1）设$c_1, c_2, \cdots, c_n$与$d_1, d_2, \cdots, d_n$是（1）任意解，则有向量

$$\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$$
$$\beta = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_n\alpha_n$$

$\alpha, \beta \in L_1$，则

$$\alpha + \beta = (c_1 + d_1)\alpha_1 + (c_2 + d_2)\alpha_2 + \cdots + (c_n + d_n)d_n$$
$$k\alpha = (kc_1)\alpha_1 + (kc_2)\alpha_2 + \cdots + (kc_n)\alpha_n$$

由于$(c_1 + d_1)$，$(c_2 + d_2), \cdots$，$(c_n + d_n)$与$kc_1, kc_2, \cdots$，kc_n仍是（1）的解，即$\alpha + \beta$，$k\alpha \in L_1$，因此L_1是$V_n(F)$的子空间。

又因（1）的系数不全为0，故系数阵的秩为1。因此解空间的维数是$n-1$，其基础解系为

$$\overline{\beta_i} = (c_{i1}, c_{i2}, \cdots, c_{in}) \quad i = 1, 2, \cdots, n-1$$

由此构造出L_1的向量组

$$\beta_i = c_{i1}\alpha_1 + c_{i2}\alpha_2 + \cdots + c_{in}\alpha_n$$

是L_1的基底，故L_1是$V_n(F)$的$n-1$维子空间。

2）仿1）可以证明当（2）的系数阵的秩为r时，L_2是$n-r$维的子空间。

10．证：设$\alpha_1, \alpha_2, \cdots, \alpha_r$是$D^{(n)}$的子空间$W$的基底，令

$$A=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_r\end{pmatrix}$$

以A为系数阵作齐次线性方程组，其解空间的基础解系为

$$\beta_1,\beta_2,\cdots,\ \beta_{n-r}$$

令

$$B=(\beta_1,\beta_2,\cdots\beta_{n-r})$$

显然有

$$AB=0$$

于是则有

$$B'A'=0$$

则以B'为系数阵的齐次线性方程组的基础解系为$\alpha_1,\alpha_2,\ \cdots,\ \alpha_r$。因此$D^{(n)}$的子空间$W$就是以$B'$为系数阵的齐次线性方程组的解空间。

11．证：由上题知$D^{(n)}$的每个子空间都是某个齐次线性方程组的解空间，而每个齐次线性方程组的解空间又都看作若干个$n-1$维子空间的交，故$D^{(n)}$的真子空间都是若干个$n-1$维子空间的交。

12．证：因$c_1\cdot c_3\neq 0$，故 $c_1\neq 0$，$c_3\neq 0$，于是有

$$\alpha=-\frac{c_2}{c_1}\beta-\frac{c_3}{c_1}\gamma$$

$$\gamma=-\frac{c_1}{c_3}\alpha-\frac{c_2}{c_3}\beta$$

由此可知向量组α,β与β,γ等价，因此有

$$L(\alpha,\beta)=L(\beta,\gamma)$$

13．证：显然（1）的解空间W_1是$n-1$维的，而且

$$e_1=(1,-1,0,\cdots,0)$$

$$e_2=(1,0,-1,\cdots,0)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$e_{n-1}=(1,0,0,\cdots,-1)$$

是W_1的基底。（2）的解空间W_2是1维的，而且

$$e_n=(1,1,\cdots,1)$$

是W_2的基底。由此进一步可以证明$e_1,e_2,\cdots,e_{n-1}$，e_n是线性无关的，于是有

$$D^{(n)}=W_1\dot{+}W_2.$$

14．证：设$\alpha_1,\alpha_2,\cdots,\alpha_r$是$W$的基底，并将该基底扩充为$V_n(F)$的基底为

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\beta_{r+1},\cdots,\beta_n$$

于是

$$W_1=L(\beta_{r+1},\beta_{r+2},\cdots,\beta_n)$$

是W的补空间，再令

$$W_2=L(\beta_{r+1}+\alpha_1,\beta_{r+2},\cdots,\beta_n)$$

易证$\alpha_1,\alpha_2,\cdots,\alpha_r$，$\beta_{r+1}+\alpha_1,\beta_{r+2},\cdots,\beta_n$是$V_n(F)$的基底。故$W_2$也是$W$的补空间。

又因为可证得$\beta_{r+1}+\alpha_1$，$\beta_{r+1},\cdots$，β_n是线性无关的，故$\beta_{r+1}+\alpha_1$不能被$\beta_{r+1},\cdots,\beta_n$线性表出，因此有

$$W_1\neq W_2$$

所以$V_n(F)$的任一非平凡子空间的补空间不止一个。

15．证：由矩阵运算可知，任一n阶矩阵都可表为对称阵与反对称阵之和，于是有

$$M_n(F)=S+T$$

又因为

$$\dim S=\frac{1}{2}n(n+1)$$

$$\dim T=\frac{1}{2}n(n-1)$$

故

$$\dim S+\dim T=\dim M_n(F)$$

因此

$$M_n(F)=S\dot{+}T.$$

第十二章　线性变换

练　习　一

1．证：1）设$\alpha=(x_1,x_2)$，$\beta=(y_1,y_2)$，

$$\sigma_1(k\alpha+l\beta)=\sigma_1(kx_1+ly_1,\ kx_2+ly_2)$$
$$=(kx_1+ly_1,\ -kx_2-ly_2)$$
$$=k(x_1,-x_2)+l(y_1,-y_2)$$
$$=k\sigma_1(\alpha)+l\sigma_1(\beta)$$

同理可证$\sigma_2,\sigma_3,\sigma_4,\sigma_5$是线性变换，其中

1）是关于X轴对称，

2）是关于Y轴对称，

3）是关于原点对称，

4）是在X轴上投影，

5）是在Y轴上投影．

2．证：令$\alpha=(x_1,x_2)$，$\beta=(y_1,y_2)$

$$\sigma(\alpha+\beta)=((x_1+y_1)\cos\theta-(x_2+y_2)\sin\theta,$$
$$(x_1+y_1)\sin\theta+(x_2+y_2)\cos\theta)$$
$$=(x_1\cos\theta-x_2\sin\theta+y_1\cos\theta-y_2\sin\theta,$$
$$x_1\sin\theta+x_2\cos\theta+y_1\sin\theta+y_2\cos\theta)$$
$$=(x_1\cos\theta-x_2\sin\theta,\ x_1\sin\theta+x_2\cos\theta)$$
$$+(y_1\cos\theta-y_2\sin\theta,y_1\sin\theta+y_2\cos\theta)$$
$$=\sigma(\alpha)\ +\sigma(\beta)$$

$$\sigma(k\alpha)=(kx_1\cos\theta-kx_2\sin\theta,kx_1\sin\theta+kx_2\cos\theta)$$

$$=k(x_1\cos\theta-x_2\sin\theta, x_1\sin\theta+x_2\cos\theta)$$

$$=k(\sigma(\alpha))$$

故σ 是线性变换.

$$\tau(\alpha+\beta)=\tau(x_1+y_1, x_2+y_2)$$

$$=(x_1+y_1+a, x_2+y_2+b)$$

$$\tau(\alpha)+\tau(\beta)=(x_1+a,\ x_2+b)+(y_1+a, y_2+b)$$

$$=(x_1+y_1+2a,\ x_2+y_2+2b)$$

因a,b 不全为 0，显然

$$\tau(\alpha+\beta)\neq\tau(\alpha)+\tau(\beta)$$

故τ 不是线性变换.

3．解：用线性变换的定义来验证

1）是线性变换.

2）因$(a_1+b_1)^2\neq a_1^2+b_1^2$,故$\sigma_2$ 不是线性变换.

3）因$\cos(a_1+b_1)\neq\cos a_1+\cos b_1$,故$\sigma_3$ 不是线性变换.

4．设X,Y是$M_n(F)$中任意二矩阵,k,l属于F,则

$$\sigma(kX+lY)=A(kX+lY)B$$

$$=A(kX)B+A(lY)B$$

$$=kAXB+lAYB$$

$$=k\sigma(X)+l\sigma(Y)$$

故σ 是线性变换.

5．证：设σ 是相似变换，故是线性变换.

反之，设ε 是V 的基底，σ 是线性变换，故有，

$$\sigma(\varepsilon)=\lambda\varepsilon$$

对任意$\alpha\in V$,必有

$$\alpha=k\varepsilon$$

$$\sigma(\alpha)=\sigma(k\varepsilon)=k(\sigma\varepsilon)=k(\lambda\varepsilon)=\lambda(k\varepsilon)=\lambda\alpha.$$

从而得证.

6．证： 1) $\sigma(kX+lY)=A(kX+lY)-(kX+lY)A$

$$=kAX+lAY-(kXA+lYA)$$

$$= k(AX - XA) + l(AY - YA)$$
$$= k\sigma(X) + l\sigma(Y)$$

故σ 是线性变换.

2) $\sigma(XY) = AXY - XYA$
$$= AXY - XAY + XAY - XYA$$
$$= (AX - XA)Y + X(AY - YA)$$
$$= \sigma(X)\cdot Y + X\sigma(Y)$$

故等式成立。

练　习　二

1．证：在几何空间$D^{(3)}$中任意一向量$\alpha = (x, y, z)$则有

$$\sigma(\alpha) = (x, -z, y)$$
$$\tau(\alpha) = (z, y, -x)$$
$$\rho(\alpha) = (-y, x, z)$$

由此可有

$$\sigma^2(\alpha) = (x, -y, -z)$$
$$\sigma^3(\alpha) = (x, z, -y)$$
$$\sigma^4(\alpha) = (x, y, z)$$

于是有

$$\sigma^4 = e$$

同理

$$\tau^2(\alpha) = (-x, y, -z)$$
$$\tau^3(\alpha) = (-z, y, x)$$
$$\tau^4(\alpha) = (x, y, z)$$
$$\tau^4 = e$$
$$\rho^2(\alpha) = (-x, -y, z)$$
$$\rho^3(\alpha) = (y, -x, z)$$
$$\rho^4(\alpha) = (x, y, z)$$

$$\rho^4 = e$$

因此有

$$\sigma^4 = \tau^4 = \rho^4 = e$$

2) $(\sigma\tau)(\alpha) = \sigma(\tau(\alpha)) = \sigma(z, y, -x) = (z, x, y)$

$(\tau\sigma)(\alpha) = \tau(\sigma(\alpha)) = \tau(x, -z, y) = (y, -z, -x)$

因此可得

$$\sigma\tau \neq \tau\sigma$$

3) $(\sigma^2\tau^2)(\alpha) = \sigma^2(\tau^2(\alpha)) = \sigma^2(-x, y, -z) = (-x, -y, z)$

$(\tau^2\sigma^2)(\alpha) = \tau^2(\sigma^2(\alpha)) = \tau^2(x, -y, -z) = (-x, -y, z)$

故有

$$\sigma^2\tau^2 = \tau^2\sigma^2$$

4) $(\sigma\tau)^2(\alpha) = (\sigma\tau)((\sigma\tau)(\alpha)) = \sigma\tau(z, x, y) = (y, z, x)$

$(\sigma^2\tau^2)(\alpha) = (-x, -y, z)$

故有

$$(\sigma\tau)^2 \neq \sigma^2\tau^2.$$

2. 对任意$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(F)$

$$(\sigma+\tau)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \sigma\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \tau\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$= \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} + \begin{pmatrix} \lambda a & 0 \\ 0 & \mu d \end{pmatrix}$$

$$= \begin{pmatrix} a+b & a-b \\ c+d & c-d \end{pmatrix} + \begin{pmatrix} \lambda a & 0 \\ 0 & \mu d \end{pmatrix}$$

$$= \begin{pmatrix} (\lambda+1)a+b & a-b \\ c+d & (\mu-1)d+c \end{pmatrix}$$

$$(\sigma\tau)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \sigma\left(\tau\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)$$

$$= \sigma\begin{pmatrix} \lambda a & 0 \\ 0 & \mu d \end{pmatrix}$$

$$=\begin{pmatrix}\lambda a & 0\\ 0 & \mu d\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$$

$$=\begin{pmatrix}\lambda a & \lambda a\\ \mu d & -\mu d\end{pmatrix}$$

3．对任意 $f(x)\in F[x]$

$$\begin{aligned}(\sigma\tau-\tau\sigma)f(x)&=(\sigma\tau)(f(x))-(\tau\sigma)(f(x))\\&=\sigma(xf(x))-\tau(f'(x))\\&=f(x)+xf'(x)-xf'(x)\\&=f(x)\end{aligned}$$

故有

$$\sigma\tau-\tau\sigma=\varepsilon_{\circ}$$

4．证：当$k=2$时，有

$$\begin{aligned}\sigma^2\tau-\tau\sigma^2&=\sigma^2\tau-\sigma\tau\sigma+\sigma\tau\sigma-\tau\sigma^2\\&=\sigma(\sigma\tau-\tau\sigma)+(\sigma\tau-\tau\sigma)\sigma\\&=\sigma+\sigma\\&=2\sigma\end{aligned}$$

命题成立。

假如$k=n$时对，即

$$\sigma^n\tau-\tau\sigma^n=n\sigma^{n-1}$$

现证$k=n+1$时也成立

$$\begin{aligned}\sigma^{n+1}\tau-\tau\sigma^{n+1}&=\sigma\sigma^n\tau-\sigma\tau\sigma^n+\sigma\tau\sigma^n-\tau\sigma^{n+1}\\&=\sigma(\sigma^n\tau-\tau\sigma^n)+(\sigma\tau-\tau\sigma)\sigma^n\\&=\sigma n\sigma^{n-1}+\sigma^n\\&=(n+1)\sigma^n\end{aligned}$$

故等式成立。

5．证： 1) $(\sigma+\tau)^2=\sigma^2+\sigma\tau+\tau\sigma+\tau^2$

$$=\sigma+\sigma\tau+\tau\sigma+\tau$$

因为

$$(\sigma+\tau)^2=\sigma+\tau$$

故有

$$\sigma\tau+\tau\sigma=\theta$$

即

$$\sigma\tau=-\tau\sigma$$

又

$$\begin{aligned}2\sigma\tau&=\sigma\tau+\sigma\tau=\sigma\tau-\tau\sigma=\sigma^2\tau-\tau\sigma^2\\&=\sigma^2\tau-\tau\sigma^2=\sigma^2\tau+\sigma\tau\sigma=\sigma(\sigma\tau+\tau\sigma)\\&=\sigma\theta=\theta\end{aligned}$$

故

$$\sigma\tau=\theta$$

2) $$\begin{aligned}(\sigma+\tau-\sigma\tau)^2&=\sigma^2+\tau\sigma-\sigma\tau\sigma+\sigma\tau+\tau^2-\sigma\tau^2-\sigma^2\tau-\tau\sigma\tau+\sigma\tau\sigma\tau\\&=\sigma+\tau\sigma-\sigma\tau\sigma+\sigma\tau+\tau-\sigma\tau-\sigma\tau-\tau\sigma\tau+\sigma\tau\sigma\tau\\&=\sigma+\sigma\tau-\sigma\tau+\sigma\tau+\tau-\sigma\tau-\sigma\tau-\sigma\tau+\sigma\tau\\&=\sigma+\tau-\sigma\tau\end{aligned}$$

故等式得证。

6．证：$[[\rho,\sigma],\tau]+[[\sigma,\tau],\rho]+[[\tau,\rho],\sigma]$

$$\begin{aligned}&=[\rho\sigma-\sigma\rho,\tau]+[\sigma\tau-\tau\sigma,\rho]+[\tau\rho-\rho\tau,\sigma]\\&=(\rho\sigma-\sigma\rho)\tau-\tau(\rho\sigma-\sigma\rho)+(\sigma\tau-\tau\sigma)\rho-\rho(\sigma\tau-\tau\sigma)\\&\quad+(\tau\rho-\rho\tau)\sigma-\sigma(\tau\rho-\rho\tau)\\&=\rho\sigma\tau-\sigma\rho\tau-\tau\rho\sigma+\tau\sigma\rho+\sigma\tau\rho-\tau\sigma\rho-\rho\sigma\tau+\rho\tau\sigma+\tau\rho\sigma\\&\quad-\rho\tau\sigma-\sigma\tau\rho+\sigma\rho\tau\\&=\theta\end{aligned}$$

故等式成立。

7．证：设 σ,τ 是可逆线性变换，令其逆变换为σ^{-1}，τ^{-1}。于是有

$$(\sigma\tau)\cdot(\tau^{-1}\sigma^{-1})=(\tau^{-1}\sigma^{-1})(\sigma\tau)=\varepsilon$$

故$\sigma\tau$是可逆线性变换。且$(\sigma\tau)^{-1}=\tau^{-1}\sigma^{-1}$。

8．证：考查等式

$$k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_n\sigma(\alpha_n)=0$$

$$\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n)=0$$

因σ 是可逆线性变换，故有σ^{-1}，用σ^{-1}去作用上式，则有

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n=0$$

因$\alpha_1,\alpha_2,\cdots,\alpha_n$是线性无关，故有

$$k_1=k_2=\cdots=k_n=0$$

因此$\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)$是线性无关的。

练　习　三

1．证：1）因$\alpha_1,\alpha_2,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$是$V_n(F)$的基底，则$\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_s),\sigma(\alpha_{s+1}),\cdots,\sigma(\alpha_n)$是$\sigma(V_n(F)$的生成组，又因$\alpha_1,\alpha_2,\cdots,\alpha_s$是$\sigma^{-1}(0)$的基底，故

$$\sigma(\alpha_i)=0\quad i=1,2,\cdots,s$$

因此$\sigma(\alpha_{s+1}),\cdots,\sigma(\alpha_n)$是$\sigma(V_n(F))$的生成组。

下面证明$\sigma(\alpha_{s+1}),\cdots,\sigma(\alpha_n)$是线性无关，设

$$k_{s+1}\sigma(\alpha_{s+1})+k_{s+2}\sigma(\alpha_{s+2})+\cdots+k_n\sigma(\alpha_n)=0$$

$$\sigma(k_{s+1}\alpha_{s+1}+k_{s+2}\alpha_{s+2}+\cdots+k_n\alpha_n)=0$$

故$k_{s+1}\alpha_{s+1}+k_{s+2}\alpha_{s+2}+\cdots+k_n\alpha_n\in\sigma^{-1}(0)$，于是有

$$k_{s+1}\alpha_{s+1}+k_{s+2}\alpha_{s+2}+\cdots+k_n\alpha_n=l_1\alpha_1+l_2\alpha_2+\cdots+l_s\alpha_s$$

即

$$l_1\alpha_1+l_2\alpha_2+\cdots+l_s\alpha_s-k_{s+1}\alpha_{s+1}-\cdots-k_n\alpha_n=0$$

因$\alpha_1,\alpha_2,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$是$V_n(F)$的基底，

故

$$k_{s+1}=k_{s+2}=\cdots=k_n=0$$

因此$\sigma(\alpha_{s+1})$，$\sigma(\alpha_{s+2}),\cdots,\sigma(\alpha_n)$，是线性无关的，从而是$\sigma(V_n(F))$的基底

2）因为

$$\dim(\sigma^{-1}(0))=s,\ \dim(\sigma(V))=n-s$$

故有

$$\dim(\sigma^{-1}(0))+\dim(\sigma(V))=n.$$

2．解：不能成立。如在$F_{n+1}[x]$中，令

$$\sigma(f(x))=f'(x)$$

显然

$$\sigma(F_{n+1}[x])=F_n[x],\quad \sigma^{-1}(0)=F$$

其中

$$\dim\sigma(F_{n+1}[x])=n$$

$$\dim(\sigma^{-1}(0))=1$$

它们的维数和显然等于$F_{n+1}[x]$的维数，但

$$\sigma(F_{n+1}[x])+\sigma^{-1}(0)\neq F_{n+1}[x].$$

3．证：必要性，设$\alpha\in\sigma^{-1}(0)$，即有

$$\sigma(\alpha)=0$$

由σ 是可逆线性变换，则有σ^{-1}，于是

$$\sigma^{-1}(\sigma(\alpha))=\sigma^{-1}(0)$$

$$\alpha=0$$

因此$\sigma^{-1}(0)=\{0\}$

充分性：由$\sigma^{-1}(0)=\{0\}$，$\dim(\sigma^{-1}(0))=0$，故

$$\dim(\sigma(V))=n$$

因此σ 是满射。由$\alpha_1\neq\alpha_2$得$\sigma(\alpha_1)\neq\sigma(\alpha_2)$

事实上若$\sigma(\alpha_1)=\sigma(\alpha_2)$则

$$\sigma(\alpha_1)-\sigma(\alpha_2)=0$$

$$\sigma(\alpha_1-\alpha_2)=0$$

由$\sigma^{-1}(0)=\{0\}$，故有$\alpha_1=\alpha_2$，与$\alpha_1\neq\alpha_2$矛盾，因此σ 是单变换，因此σ 是单满变换，即σ 为可逆变换。

4．证：因 $\sigma^{-1}(0)$与W的交 $\sigma^{-1}(0)\cap W$ 是W的子空间，设其维数为 r，W的维数是 s，令ε_1，$\varepsilon_2,\cdots,\varepsilon_r$是$\sigma^{-1}(0)\cap W$的基底，将此扩为$W$的基底，

$$\varepsilon_1,\varepsilon_2,\cdots,\quad \varepsilon_r,\varepsilon_{r+1}\cdots\varepsilon_s$$

则$\sigma(W)=L(\sigma(\varepsilon_1),\sigma(\varepsilon_2),\cdots,\sigma(\varepsilon_s))$仿 1 题可得$\sigma(\varepsilon_{r+1}),\sigma(\varepsilon_{r+2}),\cdots,$

$\sigma(\varepsilon_s)$是$\sigma(W)$的基底，故有

$$\dim(\sigma(W))=s-r=\dim W-\dim(\sigma^{-1}(0)\cap W)$$

因此

$$\dim(\sigma(W))+\dim(\sigma^{-1}(0)\cap W)=\dim W.$$

5. 证：因$\tau(V)$是$V_n(F)$的子空间，由上题有

$$\dim(\sigma\tau(V))+\dim(\sigma^{-1}(0)\cap\tau(V))=\dim(\tau(V))$$

于是

$$\dim(\sigma\tau(V))=\dim(\tau(V))-\dim(\sigma^{-1}(0)\cap\tau(V))$$

又因为对任意子空间W，都有

$$\sigma^{-1}(0)\cap W\subseteq\sigma^{-1}(0)$$

故有

$$\dim(\sigma^{-1}(0)\cap W)\leqslant\dim\sigma^{-1}(0)=n-\dim(\sigma(V))$$

因此

$$\dim(\sigma\tau(V))\geqslant\dim(\tau(V))-(n-\dim\sigma(V))$$
$$=\dim(\tau(V))+\dim(\sigma(V))-n.$$

6. 证：因为

$$\dim(\sigma\tau(V))+\dim((\sigma\tau)^{-1}(0))=n$$

故

$$\dim((\sigma\tau)^{-1}(0))=n-\dim(\sigma\tau(V))$$

再由上题知

$$\dim(\sigma\tau(V))\geqslant\dim(\sigma(V))+\dim(\tau(V))-n$$

于是

$$\dim((\sigma\tau)^{-1}(0))\leqslant n-(\dim(\sigma(V))+\dim\tau(V)-n)$$
$$=n-\dim(\sigma(V))+n-\dim(\tau(V))$$
$$=\dim\sigma^{-1}(0)+\dim\tau^{-1}(0)$$

即

$$\dim((\sigma\tau)^{-1}(0))\leqslant\dim\sigma^{-1}(0)+\dim\tau^{-1}(0).$$

练　习　四

1．证：任取$\alpha\in W$，因W是σ,τ 的不变子空间，故$\sigma(\alpha),\tau(\alpha)\in W$，则有

$$(\sigma+\tau)(\sigma)=\sigma(\alpha)+\tau(\alpha)\in W$$

故W是$\sigma+\tau$的不变子空间

$$(\sigma\cdot\tau)(\alpha)=\sigma(\tau(\alpha))\in W$$

故W是$\sigma\tau$的不变子空间。

2．证：设$\sigma_1,\sigma_2,\cdots,\sigma_r$是$W$的基底，因 σ 是可逆线性变换，故$\sigma(\sigma_1),\sigma(\alpha_2),\cdots$，$\sigma(\sigma_r)$也是$W$的基底。对任意$\alpha\in W$，则有

$$\alpha=k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_v\sigma(\alpha_r)$$

$$\sigma^{-1}(\alpha)=\sigma^{-1}(\sigma(k_1\sigma_1+k_2\alpha_2+\cdots+k_r\sigma_r))$$

$$=k_1\alpha_1+k_2\alpha_2+\cdots+k_r\sigma_r$$

因$\alpha_1,\sigma_2,\cdots,\alpha_r$是$W$的基底，故$k_1\sigma_1+k_2\alpha_2+\cdots+k_r\sigma_r$属于$W$，即 $\sigma^{-1}(\alpha)\in W$，所以W是σ^{-1}的不变子空间。

3．证：任取$\sigma\in\tau^{-1}(0)$，则有

$$\tau(\alpha)=0$$

从而有

$$\tau(\sigma(\alpha))=(\tau\sigma)(\alpha)=(\sigma\tau)(\alpha)=\sigma(\tau(\alpha))=\sigma(0)=0$$

故$\sigma(\alpha)\in\tau^{-1}(0)$，因此$\tau^{-1}(0)$是$\sigma$ 的不变子空间。

任取$\alpha\in\tau(V_n(F))$，则存在$\beta\in V_n(F)$使得

$$\tau(\beta)=\alpha$$

且$\sigma(\beta)\in V_n(F)$，则

$$\sigma(\alpha)=\sigma(\tau(\beta))=(\sigma\tau)(\beta)=(\tau\sigma)(\beta)=\tau(\sigma(\beta))\in\tau(V_n(F))$$

故$\tau(V_n(F))$是σ 的不变子空间。

4．证：

任意$\alpha\in\sigma(W)$，因$\sigma(W)\subseteq W$，故 $\sigma(\alpha)\in\sigma(W)$，因此$\sigma(W)$ 是σ的不变子空间。

又设任意 $\sigma' \in W'$，$\sigma(\sigma') \in W$，因W 是 σ 的不变子空间，故$\sigma(\sigma(\sigma')) \in W$，于是$\sigma(\sigma') \in W'$，因此$W'$是$\sigma$ 的不变子空间。

5．证：设$L(\alpha)$是$V_n(F)$的任一一维子空间，任取$\beta \in L(\alpha)$，都有$\beta = k\alpha$，又因σ 是线性变换，

故有

$$\sigma(\alpha) = \lambda\alpha$$

因此

$$\sigma(\beta) = \sigma(k\alpha) = k\sigma(\alpha) = k\lambda\alpha \in L(\alpha)$$

故$L(\alpha)$是σ 的不变子空间。

反之，设$L(\alpha)$是$V_n(F)$的任一一维子空间，且对σ 是不变的，令

$$\sigma(\alpha) = \lambda\alpha$$

去证对任意$\beta \in V_n(F)$都有

$$\sigma(\beta) = \lambda\beta.$$

当$\beta \in L(\alpha)$，即$\beta = k\alpha$则

$$\sigma(\beta) = \sigma(k\alpha) = k\sigma(\alpha) = k\lambda\sigma = \lambda(k\sigma) = \lambda\beta$$

当 $\beta \overline{\in} L(\alpha)$，则$\alpha$，$\beta$ 线性无关，则$L(\sigma+\beta)$也是σ 的一维不变子空间，于是

$$\sigma(\alpha+\beta) = l(\alpha+\beta) = l\alpha + l\beta$$

又

$$\sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta) = \lambda\alpha + \mu\beta$$

即

$$l\alpha + l\beta = \lambda\alpha + \mu\beta$$

$$(l-\lambda)\alpha + (l-\mu)\beta = 0$$

因σ，β 线性无关，则有

$$l = \mu = \lambda$$

于是有

$$\sigma(\beta) = \lambda\beta$$

所以σ 是相似变换。

6．证：必要性是显然的，下面证明充分性．对任意$\alpha\in W$，

$$\alpha=k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m$$

$$\begin{aligned}\sigma(\alpha)&=\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m)\\&=k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_m\sigma(\alpha_m)\end{aligned}$$

因$\sigma(\alpha_i)\in W$，$(i=1,2,\cdots,m)$故 $\sigma(\alpha)\in W$，所以W是σ 的不变子空间．

练 习 五

1．证：1）

$$\begin{aligned}\sigma(k\alpha+l\beta)&=\sigma(k(x_1,x_2,x_3)+l(y_1,y_2,y_3))\\&=\sigma(kx_1+ly_1,kx_2+ly_2,kx_3+ly_3)\\&=(2(kx_1+ly_1)-(kx_2+ly_2),\\&\quad k_2x_2+ly_2+kx_3+ly_3,\ kx_1+ly_1)\\&=(2kx_1-kx_2+2ly_1-ly_2,\\&\quad kx_2+kx_3+ly_2+ly_3,kx_1+ly_1)\\&=k(2x_1-x_2,\ x_2+x_3,x_1)\\&\quad+l(2y_1-y_2,\ y_2+y_3,y_1)\\&=k\sigma(\alpha)+l\sigma(\beta)\end{aligned}$$

故σ 是线性变换．

2）因为

$$\sigma(\varepsilon_1)=(2,0,1,)=2\varepsilon_1+\varepsilon_3$$

$$\sigma(\varepsilon_2)=(-1,1,0)=-\varepsilon_1+\varepsilon_2$$

$$\sigma(\varepsilon_3)=(0,1,\ 0)=\varepsilon_2$$

故σ 在$\varepsilon_1,\varepsilon_2,\varepsilon_3$ 上的表示矩阵为

$$\begin{pmatrix}2&-1&0\\0&1&1\\1&0&0\end{pmatrix}.$$

2．解：1）因σ 在$\varepsilon_1,\varepsilon_2,\varepsilon_3$，上的表示矩阵为

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

故有

$$\sigma(\varepsilon_1) = a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + a_{31}\varepsilon_3$$

$$\sigma(\varepsilon_2) = a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + a_{32}\varepsilon_3$$

$$\sigma(\varepsilon_3) = a_{13}\varepsilon_1 + a_{23}\varepsilon_2 + a_{33}\varepsilon_3$$

于是有

$$\sigma(\varepsilon_3) = a_{33}\varepsilon_3 + a_{23}\varepsilon_2 + a_{13}\varepsilon_1$$

$$\sigma(\varepsilon_2) = a_{32}\varepsilon_3 + a_{22}\varepsilon_2 + a_{12}\varepsilon_1$$

$$\sigma(\varepsilon_1) = a_{31}\varepsilon_3 + a_{21}\varepsilon_2 + a_{11}\varepsilon_1$$

故σ 在ε_3，ε_2，ε_1上的表示矩阵为

$$\begin{pmatrix} a_{33} & a_{32} & a_{31} \\ a_{23} & a_{22} & a_{21} \\ a_{13} & a_{12} & a_{11} \end{pmatrix}$$

同理可求得

2） σ 在$\varepsilon_1, k\varepsilon_2, \varepsilon_3$上的表示矩阵为

$$\begin{pmatrix} a_{11} & ka_{12} & a_{13} \\ a_{21} & ka_{22} & a_{23} \\ a_{31} & ka_{32} & a_{33} \end{pmatrix}$$

3） σ 在$\varepsilon_1+\varepsilon_2$，$\varepsilon_2, \varepsilon_3$上的表示矩阵为

$$\begin{pmatrix} a_{11}+a_{12} & a_{12} & a_{13} \\ a_{21}+a_{22} & a_{22} & a_{23} \\ a_{31}+a_{32} & a_{32} & a_{33} \end{pmatrix}$$

3．解：1）因为

$\sigma(1) = 0$

$\sigma(1+x) = 1$

$\sigma(1+x^2) = 2x = -2+2(1+x)$

……

$$\sigma(1+x^{n-1})=(n-1)x^{n-2}=-(n-1)+(n-1)(1+x^{n-2})$$

故σ 在基底 1，$1+x, 1+x^2, \cdots, 1+x^{n-1}$上的表示矩阵为

$$\begin{pmatrix} 0 & 1 & -2 & \cdots & -(n-1) \\ 0 & 0 & 2 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & n-1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

2）因为

$$\sigma(1)=0$$

$$\sigma(x-c)=1$$

$$\sigma\left(\frac{(x-c)^2}{2!}\right)=x-c$$

$$\sigma\left(\frac{(x-c)^3}{3!}\right)=\frac{(x-c)^2}{2!}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sigma\frac{(x-c)^{n-1}}{(n-1)!}=\frac{(x-c)^{n-2}}{(n-2)!}$$

故σ 在 1，$x-c$，$\frac{(x-c)^2}{2!}, \cdots, \frac{(x-c)^{n-1}}{(n-1)!}$上的表示矩阵为

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

4．任取$\alpha=(x_1, x_2, x_3)\in D^{(3)}$，则$\alpha$ 在基底ε_1，$\varepsilon_2, \varepsilon_3$上的坐标为$(x_1-x_2)$，$(x_2-x_3)$，$x_3$，故

$$\sigma(x_1, x_2, x_3)=(\varepsilon_1, \varepsilon_2, \varepsilon_3)\begin{pmatrix} 1 & -1 & 2 \\ -1 & 0 & -1 \\ 1 & 2 & 2 \end{pmatrix}\begin{pmatrix} x_1-x_2 \\ x_2-x_3 \\ x_3 \end{pmatrix}$$

$$=(x_1-2x_2+3x_3)\varepsilon_1+(-x_1+x_2-x_3)\varepsilon_2+(x_1+x_2)\varepsilon_3$$

$$=(x_1+2x_3, 2x_2-x_3, x_1+x_2)$$

因此

$$\sigma(1,2,3)=(7,1,3)$$

即α 在σ 下的象是（7，1，3）。

5．解：因为

$$\sigma(\eta_1)=(-5,0,3)=2\eta_1-\eta_2-\eta_3$$

$$\sigma(\eta_2)=(0,-1,6)=3\eta_1-\eta_3$$

$$\sigma(\eta_3)=(-5,-1,9)=5\eta_1-\eta_2$$

故σ 在η_1,η_2,η_3，上的表示矩阵为

$$A=\begin{pmatrix}2&3&5\\-1&0&-1\\-1&-1&0\end{pmatrix}.$$

练　习　六

1．由题设可知由$\{\alpha_i\}$到$\{\beta_i\}$的过渡阵为

$$T=\begin{pmatrix}-1&1&-1\\0&1&-1\\1&0&1\end{pmatrix}, \text{而} T^{-1}=\begin{pmatrix}-1&1&0\\1&0&1\\1&-1&1\end{pmatrix}$$

于是 σ 在$\{\beta_i\}$上的表示矩阵为

$$B=T^{-1}AT=\begin{pmatrix}-1&1&0\\1&0&1\\1&-1&1\end{pmatrix}\begin{pmatrix}1&0&1\\1&1&0\\-1&2&0\end{pmatrix}\begin{pmatrix}-1&1&-1\\0&1&-1\\1&0&1\end{pmatrix}$$

$$=\begin{pmatrix}-1&1&-2\\1&2&-1\\2&0&1\end{pmatrix}$$

2．1）容易算得

$$\eta_1=-2\varepsilon_1+\varepsilon_2+\varepsilon_3,$$

$$\eta_2=-\frac{3}{2}\varepsilon_1+\frac{3}{2}\varepsilon_2+\frac{1}{2}\varepsilon_3,$$

$$\eta_3 = \frac{3}{2}\varepsilon_1 + \frac{3}{2}\varepsilon_2 - \frac{5}{2}\varepsilon_3.$$

所以从$\{\varepsilon_i\}$到$\{\eta_i\}$的过渡阵为

$$A = \begin{pmatrix} -2 & -\frac{3}{2} & \frac{3}{2} \\ 1 & \frac{3}{2} & \frac{3}{2} \\ 1 & \frac{1}{2} & -\frac{5}{2} \end{pmatrix}$$

2） 由 $\sigma(\varepsilon_i) = \eta_i$，可见$\sigma$ 在$\{\varepsilon_i\}$上的表示矩阵就是A。

3） 由$\sigma(\eta_1\eta_2\eta_3) = \sigma(\varepsilon_1\varepsilon_2\varepsilon_3)A = (\eta_1\eta_2\eta_3)A$。于是$\sigma$ 在$\{\eta_i\}$上的表示矩阵还是A。

习 题 十二

1．证：将$\alpha_1, \alpha_2, \cdots, \alpha_m$扩为$V_n(F)$的基底：$\alpha_1, \alpha_2, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$；将$\beta_1, \beta_2, \cdots \beta_m$，再添加$n-m$个向量，构成一个含有$n$个向量的向量组$\beta_1, \beta_2, \cdots, \beta_m, \beta_{m+1}, \cdots, \beta_n$，

对$V_n(F)$中任意向量α，都有

$$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$

令$V_n(F)$ 中一变换 σ，使得

$$\sigma(\alpha) = k_1\beta_1 + k_2\beta_2 + \cdots + k_n\beta_n$$

下面首先证明σ 是线性变换，设α，β 是$V_n(F)$ 中任意两个向量，有

$$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$

$$\beta = l_1\alpha_1 + l_2\alpha_2 + \cdots + l_n\alpha_n$$

$$\begin{aligned} \sigma(\alpha+\beta) &= \sigma((k_1+l_1)\alpha_1 + (k_2+l_2)\alpha_2 + \cdots + (k_n+l_n)\alpha_n) \\ &= (k_1+l_1)\beta_1 + (k_2+l_2)\beta_2 + \cdots + (k_n+l_n)\beta_n \\ &= k_1\beta_1 + k_2\beta_2 + \cdots + k_n\beta_n + l_1\beta_1 + l_2\beta_2 + \cdots + l_n\beta_n \\ &= \sigma(\alpha) + \sigma(\beta) \end{aligned}$$

$$\sigma(k\alpha)=\sigma(kk_1\alpha+kk_2\alpha_2+\cdots+kk_n\alpha_n)$$
$$=kk_1\beta_1+kk_2\beta_2+\cdots+kk_n\beta_n$$
$$=k\sigma(\alpha)$$

故σ 是线性变换。

而且由σ 的规定立即可得

$$\sigma(\alpha_i)=\beta_i.$$

2．解：1）因为

$$\sigma(\varepsilon_0)=1-1=0$$

$$\sigma(\varepsilon_1)=x+1-x=1=\varepsilon_0$$

$$\vdots$$

$$\sigma(\varepsilon_{n-1})=\frac{(x+1)(x)\cdots(x-(n-3))}{(n-1)!}$$

$$-\frac{x(x-1)\cdots(x-(n-2))}{(n-1)!}$$

$$=\frac{x(x-1)\cdots(x-(x-3))}{(n-1)!}$$

$$\cdot[(x+1)-(x-(n-2))]$$

$$=\frac{x(x-1)\cdots(x-(n-3))}{(n-1)!}\cdot(n-1)$$

$$=\frac{x(x-1)\cdots(x-(n-3))}{(n-2)!}$$

$$=\varepsilon_{n-2}$$

由此可得σ 的表示矩阵为

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

2）因为

$$\sigma(\varepsilon_1)=a\varepsilon_1-b\varepsilon_2$$

$$\sigma(\varepsilon_2) = b\varepsilon_1 + a\varepsilon_2$$

$$\sigma(\varepsilon_3) = \varepsilon_1 + a\varepsilon_3 - b\varepsilon_4$$

$$\sigma(\varepsilon_4) = \varepsilon_2 + b\varepsilon_3 + a\varepsilon_4$$

$$\sigma(\varepsilon_5) = \varepsilon_3 + a\varepsilon_5 - b\varepsilon_6$$

$$\sigma(\varepsilon_6) = \varepsilon_4 + b\varepsilon_6 + a\varepsilon_5$$

故σ 的表示矩阵为

$$\begin{pmatrix} a & b & 1 & 0 & 0 & 0 \\ -b & a & 0 & 1 & 0 & 0 \\ 0 & 0 & a & b & 1 & 0 \\ 0 & 0 & -b & a & 0 & 1 \\ 0 & 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & -b & a \end{pmatrix}.$$

3．解：因为

$$\sigma_1(E_{11}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} = aE_{11} + cE_{21}$$

$$\sigma_1(E_{12}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a \\ 0 & c \end{pmatrix} = aE_{12} + cE_{22}$$

$$\sigma_1(E_{21}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} b & 0 \\ d & 0 \end{pmatrix} = bE_{11} + dE_{21}$$

$$\sigma_1(E_{22}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & b \\ 0 & d \end{pmatrix} = bE_{12} + dE_{22}$$

故σ_1的表示矩阵为

$$A_1 = \begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix},$$

又因为

$$\sigma_2(E_{11}) = aE_{11} + bE_{12}$$

$$\sigma_2(E_{12}) = cE_{11} + dE_{12}$$

$$\sigma_2(E_{21}) = aE_{21} + bE_{22}$$

$$\sigma_2(E_{22}) = cE_{21} + dE_{22}$$

故σ_2的表示矩阵为

$$A_2=\begin{pmatrix}a&c&0&0\\b&d&0&0\\0&0&a&c\\0&0&b&d\end{pmatrix}$$

又因为

$$\sigma_3(E_{11})=a^2E_{11}+abE_{12}+acE_{21}+bcE_{22}$$
$$\sigma_3(E_{12})=acE_{11}+adE_{12}+c^2E_{21}+cdE_{22}$$
$$\sigma_3(E_{21})=abE_{11}+b^2E_{12}+adE_{21}+bdE_{22}$$
$$\sigma_3(E_{22})=bcE_{11}+bdE_{12}+cdE_{21}+d^2E_{22}$$

故σ_3的表示矩阵为

$$A_3=\begin{pmatrix}a^2&ac&ab&bc\\ab&ad&b^2&bd\\ac&c^2&ad&cd\\bc&cd&bd&d^2\end{pmatrix}.$$

4．证：因为当基底选定之后，线性变换与其表示矩阵之间是一一对应的．由矩阵的运算知，与一切n阶方阵可换的矩阵是纯量矩阵，从而可知与一切线性变换可换的线性变换是相似变换．

5．证：设σ在基底$\alpha_1,\alpha_2,\cdots,\alpha_n$上的表示阵矩为$A=(a_{ij})$，下而证明$A$是纯量矩阵．

设X为非奇异矩阵，有

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)X$$

则$\beta_1,\beta_2,\cdots,\beta_n$也是$V_n(F)$的基底，$\sigma$在$\{\beta_i\}$上的表示矩阵，由题设知仍为$A$，并且有

$$A=X^{-1}AX$$

从而有

$$XA=AX$$

由此而得A与一切非奇异矩阵可换．如果取

$$X_1=\begin{pmatrix}1&&&\\&2&&\\&&\ddots&\\&&&n\end{pmatrix}$$

则由

$$X_1A=AX_1$$

知当$i\neq j$时，$a_{ij}=0$

再取

$$X_2=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1\\1&0&0&\cdots&0\end{pmatrix}$$

由

$$X_2A=AX_2$$

知

$$a_{11}=a_{22}=\cdots=a_{nn}$$

从而A是纯量矩阵，故σ 是相似变换。

6．证：假如$\alpha_1+\alpha_2$ 是σ 的特征向量，于是有

$$\sigma(\alpha_1+\alpha_2)=\lambda(\alpha_1+\alpha_2)=\lambda\alpha_1+\lambda\alpha_2$$

$$\sigma(\alpha_1+\alpha_2)=\sigma(\alpha_1)+\sigma(\alpha_2)=\lambda_1\alpha_1+\lambda_2\alpha_2$$

于是有

$$(\lambda-\lambda_1)\alpha_1+(\lambda-\lambda_2)\alpha_2=0$$

由于α_1,α_2是属于不同特征根的特征向量，故线性无关，于是有

$$\lambda-\lambda_1=0$$

$$\lambda-\lambda_2=0$$

即

$$\lambda_1=\lambda=\lambda_2$$

与$\lambda_1\neq\lambda_2$矛盾，因此$\alpha_1+\alpha_2$不是σ 的特征向量。

7．证：必要性是显然的，下面证明充分性，

设$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$ 的一个基底，并设

$$\sigma(\alpha_i)=\lambda_i\alpha_i\qquad i=1,2,\cdots,n$$

由于$\alpha_i+\alpha_j$仍是σ 的特征向量，由上题有

$$\lambda_i=\lambda_j$$

故

$$\sigma(\alpha_i)=\lambda\alpha_i$$

从而对任意向量α，都有

$$\alpha=k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n$$

$$\begin{aligned}\sigma(\alpha)&=\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n)\\&=(k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_n\sigma(\alpha_n))\\&=\lambda(k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n)\\&=\lambda\alpha\end{aligned}$$

因此σ 是相似变换。

8．在$F_n[x]$中，取基底 $1, x, \frac{x^2}{2!}, \cdots, \frac{x^{n-1}}{(n-1)!}$，

则σ 在此基底上的阵为

$$A=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{pmatrix}$$

从而有

$$|\lambda I-A|=\lambda^n$$

故σ 的特征根$\lambda=0$，由于

$$(0I-A)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=0$$

的解空间的基础解系为$(1,0,\cdots,0)$故$S_\sigma(0)$ 的维数是1，小于线性空间的维数，从而在任何基底上的阵都不能与对角阵相似。

9．证：1）设$\alpha\in S_\sigma(\lambda_0)$，则有

$$\sigma(\alpha)=\lambda_0\alpha$$

从而

$$\begin{aligned}\sigma(\tau(\alpha))&=(\sigma\tau)(\alpha)=(\tau\sigma)(\alpha)=\tau(\sigma(\alpha))=\tau(\lambda_0\alpha)\\&=\lambda_0(\tau(\alpha))\end{aligned}$$

故$\tau(\alpha)\in S_\sigma(\lambda_0)$，因此$S_\sigma(\lambda_0)$是 τ 的不变子空间。

2）由$S_\sigma(\lambda_0)$是 τ 的不变子空间，在复数域上τ 在$S_\sigma(\lambda_0)$ 中必有特征根 μ，使得$0\neq\alpha\in S_\sigma(\lambda_0)$，有

$$\tau(\alpha)=\mu\alpha$$

又因为$S_\sigma(\lambda_0)$是σ 的特征子空间，因此有

$$\sigma(\alpha)=\lambda_0\alpha$$

故α 是σ，τ的共公特征向量。

10．证：1）设

$$(\sigma(\varepsilon_1),\sigma(\varepsilon_2),\cdots,\sigma(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}\lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda\end{pmatrix}$$

即

$$\sigma(\varepsilon_1)=\lambda\varepsilon_1+\varepsilon_2 \qquad (1)$$

$$\sigma(\varepsilon_2)=\lambda\varepsilon_2+\varepsilon_3 \qquad (2)$$

$$\cdots\cdots \qquad \cdots\cdots$$

$$\sigma(\varepsilon_{n-1})=\lambda\varepsilon_{n-1}+\varepsilon_n \qquad (n-1)$$

$$\sigma(\varepsilon_n)=\lambda\varepsilon_n \qquad (n)$$

令W是包含ε_1的不变子空间，则$\sigma(\varepsilon_1)\in W$，由（1）有

$$\varepsilon_2=\sigma(\varepsilon_1)-\lambda\varepsilon_1$$

故$\varepsilon_2\in W$．于是$\sigma(\varepsilon_2)$，$\lambda\varepsilon_2\in W$，故$\varepsilon_3\in W$，如此推下去$\varepsilon_n\in W$，即$\varepsilon_1$，$\varepsilon_2$，…，$\varepsilon_n\in W$，因此

$$W=V$$

2）设W是任一非零的不变子空间，令α 是W中不为 0 的向量

$$\alpha=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n$$

不妨设$k_1\neq0$，于是

$$\begin{aligned}\sigma(\alpha)&=k_1\sigma(\varepsilon_1)+k_2\sigma(\varepsilon_2)+\cdots+k_n\sigma(\varepsilon_n)\\&=k_1(\lambda\varepsilon_1+\varepsilon_2)+k_2(\lambda\varepsilon_2+\varepsilon_3)+\cdots+k_{n-1}(\lambda\varepsilon_{n-1}+\varepsilon_n)\\&\quad+k_n\lambda\varepsilon_n\end{aligned}$$

$$=\lambda(k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_n\varepsilon_n)+k_1\varepsilon_2+k_2\varepsilon_3+\cdots+k_{n-1}\varepsilon_n$$

令$k_1\varepsilon_2+k_2\varepsilon_3+\cdots+k_{n-1}\varepsilon_n=\beta$，于是有

$$\sigma(\alpha)=\lambda\alpha+\beta$$

由于$\sigma(\alpha)$，$\lambda\alpha\in W$，故$\beta\in W$，再求 $\sigma(\beta)$可有$k_1\varepsilon_3+k_2\varepsilon_4+\cdots+k_{n-2}\varepsilon_n\in W$，依次做下去，则有

$$k_1\varepsilon_n\in W,$$

从而有$\varepsilon_n\in W$。

3）设W_1，W_2是V的任意两个非平凡子空间，由2）有$\varepsilon_n\in W_1$，$\varepsilon_n\in W_2$，故

$$W_1\cap W_2\neq\{0\}$$

因此V不能分解成两个非平凡子空间的直和。

11．证：在$V_n(F)$中取定基底之后线性变换与n阶方阵是一一对应的且保持运算，故 $M_n(F)$ 与全体线性变换所组成的线性空间同构，由于 $M_n(F)$是n^2维，故全体线性变换组成的线性空间也是 n^2 维的。

12．证：1）由上题知，全体线性变换组成的线性空间的维数是n^2，故

$$\sigma^{n^2},\sigma^{n^2-1},\cdots,\sigma,\varepsilon$$

必线性相关，即存在不全为0的数$a_{n^2},a_{n^2-1},\cdots,a_1,a_0$使得

$$a_{n^2}\sigma^{n^2}+a_{n^2-1}\sigma^{n^2-1}+\cdots+a_1\sigma+a_0\varepsilon=0$$

取

$$f(x)=a_{n^2}x^{n^2}+a_{n^2-1}x^{n^2-1}+\cdots+a_1x+a_0$$

则$f(x)$的次数$\leqslant n^2$，而且有

$$f(\sigma)=0$$

2）因为

$$d(x)=(f(x),g(x))$$

故存在$u(x)$，$v(x)$使得

$$u(x)f(x)+v(x)g(x)=d(x)$$

故 $$u(\sigma)f(\sigma)+v(\sigma)g(\sigma)=d(\sigma)$$

因$f(\sigma)=0$，$g(\sigma)=0$，所以

$$d(\sigma)=0$$

3）必要性：由1）可有

$$a_{n^2}x^{n^2}+a_{n^2-1}x^{n^2-1}+\cdots+a_1x+a_0$$

是以 σ 为根的多项式，其中a_i 不全为0，若 $a_0\neq0$ 则多项式为所求。若$a_0=0$，令a_j 是不为0的系数中下标最小的那一个。即多项式为

$$a_{n^2}x^{n^2}+a_{n^2-1}x^{n^2-1}+\cdots+a_jx^j$$

且以σ 为根，有

$$a_{n^2}(\sigma)^{n^2}+a_{n^2-1}(\sigma)^{n^2-1}+\cdots+a_j(\sigma)^j=0$$

因σ 是可逆的，有σ^{-1}且$(\sigma^{-1})^j=(\sigma^j)^{-1}$也存在。故用$(\sigma^j)^{-1}$去作用上式的两端，则有

$$a_{n^2}\sigma^{n^2-j}+a_{n^2-1}\ \ \sigma^{n^2-j-1}\ \ +\cdots+a_{j+1}\sigma+a_j\varepsilon=0$$

故

$$g(x)=a_{n^2}x^{n^2-j}+a_{n^2-1}x^{n^2-j-1}+\cdots+\ +a_{j+1}x+a_j$$

即为所求。

充分性：设有常数项不为0的多项式

$$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0\qquad a_0\neq0$$

使得

$$f(\sigma)=0$$

于是有

$$a_m\sigma^m+a_{m-1}\sigma^{m-1}+\cdots+a_1\sigma=-a_0\varepsilon$$

因$a_0\neq0$，故有

$$\left(-\frac{a_m}{a_0}\sigma^{m-1}-\frac{a_{m-1}}{a_0}\sigma^{m-2}-\cdots-\frac{a_1}{a_0}\varepsilon\right)\sigma=\varepsilon$$

故σ 是可逆的。

13．证：1）设σ 的表示矩阵是A，由于σ 是可逆的，则A是非奇异的。且A的特征根就是σ 的特征根，因A是非奇异时，特征根不为0，故σ 的特征根也不为0。

2）设λ是σ的特征根，α是σ属于特征根λ的特征向量，于是有

$$\sigma(\alpha)=\lambda\alpha$$

因σ是可逆线性变换，故用σ^{-1}去作用上式，则有

$$\alpha=\lambda\sigma^{-1}(\alpha)$$

由1）知$\lambda\neq0$。故有

$$\sigma^{-1}(\alpha)=\frac{1}{\lambda}\alpha$$

所以λ^{-1}是σ^{-1}的特征根。

14．证：由特征多项式的性质知σ的特征根之积等于$|A|$，于是$|A|=0$当且仅当至少有一个特征根为0，也就是σ以0为特征根。

15．证：设$V_{ij}=\{\alpha\in V|\ \sigma_i(\alpha)=\sigma_j(\alpha)\}$显然$0\in V_{ij}$故$V_{ij}$非空。又由于$\sigma_i\neq\sigma_j$，则必存在$\beta$使得

$$\sigma_i(\beta)\neq\sigma_j(\beta)$$

即V_{ij}是V的真子集。下面证明V_{ij}是V的子空间。

对任意$\alpha,\beta\in V_{ij}$，则

$$\sigma_i(\alpha+\beta)=\sigma_i(\alpha)+\sigma_i(\beta)=\sigma_j(\alpha)+\sigma_j(\beta)=\sigma_j(\alpha+\beta)$$

故$\alpha+\beta\in V_{ij}$

$$\sigma_i(k\alpha)=k\sigma_i(\alpha)=k\sigma_j(\alpha)=\sigma_j(k\alpha)$$

故$k\alpha\in V_{ij}$，则V_{ij}是V的真子空间。

如果当V_{ij}都是V的非平凡的子空间，则由线性空间一章的学习中知必有$\alpha\in V_{ij}$，使

$$\sigma_i(\alpha)\neq\sigma_j(\alpha)$$

即存在向量α使$\sigma_1(\alpha)$，$\sigma_2(\alpha)$，…，$\sigma_s(\alpha)$两两不同。

如果V_{ij}有平凡子空间，那也只能是零空间，对于零空间只要$\alpha\neq0$，则都有

$$\sigma_i(\alpha)\neq\sigma_j(\alpha)$$

因此去掉这样子空间，只须考虑非平凡子空间则问题同前，从而得

证。

16．证：1）必要性：任取$\alpha\in V$，则$\tau(\alpha)\in\tau(V)$，又因$\sigma(V)=\tau(V)$故$\tau(\alpha)\in\sigma(V)$，则存在$\beta\in V$，使得$\tau(\alpha)=\sigma(\beta)$，于是

$$(\sigma\tau)(\alpha)=\sigma(\tau(\alpha))=\sigma(\sigma(\beta))=\sigma^2(\beta)=\sigma(\beta)=\tau(\alpha)$$

因此

$$\sigma\tau=\tau$$

同理可证

$$\tau\sigma=\sigma$$

充分性：任取$\sigma(\alpha)\in\sigma(V)\subseteq V$，因为$\sigma\tau=\tau$，$\tau\sigma=\sigma$，于是有

$$\sigma(\alpha)=(\tau\sigma)(\alpha)=\tau(\sigma(\alpha))\in\tau(V)$$

因此$\sigma(V)\subseteq\tau(V)$，同理可证$\tau(V)\subseteq\sigma(V)$，所以

$$\sigma(V)=\tau(V)$$

2）必要性，对任意$\beta\in V$，作向量$\beta-\sigma(\beta)$，则

$$\sigma(\beta-\sigma(\beta))=\sigma(\beta)-\sigma^2(\beta)=\sigma(\beta))-\sigma(\beta)=0$$

故$\beta-\sigma(\beta)\in\sigma^{-1}(0)$。因为

$$\sigma^{-1}(0)=\tau^{-1}(0)$$

所以$\beta-\sigma(\beta)\in\tau^{-1}(0)$即

$$\tau(\beta-\sigma(\beta))=\tau(\beta)-(\tau\sigma)(\beta)=0$$

因此

$$\tau(\beta)-(\tau\sigma)(\beta)=0$$

故

$$\tau(\beta)=(\tau\sigma)(\beta)$$

$$\tau=\tau\sigma$$

又因为

$$\tau(\beta-\tau(\beta))=\tau(\beta)-\tau^2(\beta)=\tau(\beta)-\tau(\beta)=0$$

故$\beta-\tau(\beta)\in\tau^{-1}(0)$，于是$\beta-\tau(\beta)\in\sigma^{-1}(0)$，于是

$$\sigma(\beta-\tau(\beta))=\sigma(\beta)-\sigma\tau(\beta)=0$$

$$\sigma(\beta)=\sigma\tau(\beta)$$

所以

$$\sigma=\sigma\tau$$

总之即有

$$\tau\sigma=\tau,\quad \sigma\tau=\sigma.$$

充分性：任取$\alpha\in\sigma^{-1}(0)$，由 $\tau\sigma=\tau$有

$$\tau(\alpha)=(\tau\sigma)(\alpha)=\tau(\sigma(\alpha))=\tau(0)=0$$

所以$\alpha\in\tau^{-1}(0)$，于是 $\sigma^{-1}(0)\subseteq\tau^{-1}(0)$

任取 $\beta\in\tau^{-1}(0)$，$\sigma\tau=\sigma$有

$$\sigma(\beta)=\sigma\tau(\beta)=\sigma(0)=0$$

故$\beta\in\sigma^{-1}(0)$。于是$\tau^{-1}(0)\subseteq\sigma^{-1}(0)$，综合便得

$$\sigma^{-1}(0)=\tau^{-1}(0).$$

17．解：因 $\alpha_1,\alpha_2,\cdots,\alpha_n$是线性无关的，故是$V_n(F)$的基底，因此$A$是非奇异的则由（1）便有

$$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A^{-1}$$

由（3）便有

$$\sigma(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\beta_1,\beta_2,\cdots,\beta_n)$$

于是

$$\begin{aligned}\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)&=\sigma(\alpha_1,\alpha_2,\cdots,\alpha_n)A^{-1}\\&=(\beta_1\beta_2\cdots\beta_n)A^{-1}\\&=(\varepsilon_1\varepsilon_2\cdots\varepsilon_n)BA^{-1}\end{aligned}$$

故σ 在$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$上的表示矩阵为$BA^{-1}$。

18．证：必要性：因为σ 是可逆变换，故σ 必是满变换，即有

$$\sigma(V_n(F))=V_n(F)$$

对任意$\alpha\in V_n(F)$，都存在$\beta\in V_n(F)$，使得

$$\sigma(\beta)=\alpha$$

由于

$$V_n(F)=W_1\dot{+}W_2$$

故有

$$\beta=\beta_1+\beta_2\quad \beta_1\in W_1,\ \beta_2\in W_2.$$

于是

$$\alpha=\sigma(\beta)=\sigma(\beta_1+\beta_2)=\sigma(\beta_1)+\sigma(\beta_2),\sigma(\beta_i)\in\sigma W_i\quad i=1,2,$$

因此有

$$V_n(F)=\sigma W_1+\sigma W_2$$

任取$\alpha\in\sigma W_1\cap\sigma W_2$，由 $\alpha\in\sigma W_1$，则有$\beta_1\in W_1$，使得

$$\sigma(\beta_1)=\alpha\qquad\qquad(1)$$

又$\alpha\in\sigma W_2$，故存在$\beta_2\in W_2$，使得

$$\sigma(\beta_2)=\alpha$$

于是有

$$\sigma(\beta_1)=\sigma(\beta_2)$$

由σ 是可逆变换，用σ^{-1}去作用上式则有

$$\beta_1=\beta_2$$

于是

$$\beta_1\in W_1\cap W_2=\{0\}$$

即 $\beta_1=0$，则由（1）可有 $\alpha=0$，故

$$\sigma W_1\cap\sigma W_2=\{0\}$$

因此

$$V_n(F)=\sigma W_1\oplus\sigma W_2$$

充分性：任取$\alpha\in V_n(F)$，由

$$V_n(F)=\sigma W_1\oplus\sigma W_2$$

有

$$\alpha=\sigma(\alpha_1)+\sigma(\alpha_2)=\sigma(\alpha_1+\alpha_2)\in\sigma(V_n(F))\quad\alpha_i\in W_i,$$

于是

$$V_n(F)\subseteq\sigma(V_n(F))$$

故

$$V_n(F)=\sigma(V_n(F))$$

由此可知σ 是满变换，因而σ 是可逆线性变换。

19．证：因为σ 与S中每个线性变换可交换于是有$\sigma^{-1}(0)$是S的不变子空间，由于S是不可约的，则$\sigma^{-1}(0)=V_n(F)$ 或 $\sigma^{-1}(0)=0$。

当$\sigma^{-1}(0)=V$，则对任意$\alpha\in V_n(F)$都有

$$\sigma(\alpha)=0$$

故σ 是零变换，

当$\sigma^{-1}(0)=\{0\}$,则σ 是单变换，从而是可逆线性变换。

20．证：当σ 是可逆线性变换，显然有

$$V=\sigma(V)=\sigma^2(V)=\cdots$$

即$k=1\leqslant n$。

当σ 不是可逆线性变换，则有

$$V\supseteq\sigma(V)\supseteq\sigma^2(V)\supseteq\cdots$$

即

$$n=\dim V\geqslant\dim\sigma(V)\geqslant\dim\sigma^2(V)\geqslant\cdots$$

由于n 是有限的，所以维数逐次减少的现象不能永远延续下去，必存在$k\leqslant n$使得

$$\sigma^k(V)=\sigma^{k+1}(V)$$

可以证明对任意正整数 m，都有

$$\sigma^{k+m}(V)=\sigma^k(V).$$

用归纳法证明：当$m=1$时显然成立。假设$m-1$ 时也成立，下面证明m 时也成立。

$$\begin{aligned}\sigma^{k+m}(V)&=\sigma\cdot(\sigma^{k+(m-1)}(V))\\&=\sigma(\sigma^k(V))\\&=\sigma^{k+1}(V)\\&=\sigma^k(V)\end{aligned}$$

故任意的m 都有

$$\sigma^{k+m}(V)=\sigma^k(V)$$

因此有$k\leqslant n$使得

$$\sigma^k(V)=\sigma^{k+1}(V)=\cdots.$$

第十三章　欧氏空间及其线性变换

练　习　一

1．解：1）不构成欧氏空间．如$\alpha=(1,0,0)\in D^{(3)}$ 且是一非零向量，但

$$(\alpha,\alpha)=1\times 0+0\times 0+0\times 1=0$$

与欧氏空间定义的第4条矛盾，故不是欧氏空间．

2）是欧氏空间．从规定的数积可直接看出

$$(\alpha,\beta)=(\beta,\alpha)$$

而且容易验证

$$(\alpha+\beta,\gamma)=(\alpha,\gamma)+(\beta,\gamma)$$

$$(k\alpha,\gamma)=k(\alpha,\gamma)$$

下面我们验证它满足第4条．

$$(\alpha,\alpha)=a_1^2+a_2^2+a_3^2+a_1a_3$$

$$=\left(a_1+\frac{1}{2}a_3\right)^2+a_2^2+\frac{3}{4}a_3^2$$

当$\alpha\neq 0$时，即 a_1,a_2,a_3不同时为零时，显然有

$$(\alpha,\alpha)>0$$

也只有a_1,a_2,a_3全为零，即$\alpha=0$时，才有

$$(\alpha,\alpha)=0$$

反之亦然，故满足第4条．因此$D^{(3)}$是欧氏空间．

2．证：1）因为

$$|\alpha+\beta|^2+|\alpha-\beta|^2=(\alpha+\beta,\alpha+\beta)+(\alpha-\beta,\alpha-\beta)$$

$$= (\alpha\cdot\alpha) + 2(\alpha\cdot\beta) + (\beta\cdot\beta) + ((\alpha\cdot\alpha)$$
$$- 2(\alpha\cdot\beta) + (\beta\cdot\beta))$$
$$= 2(\alpha\cdot\alpha) + 2(\beta\cdot\beta)$$
$$= 2|\alpha|^2 + 2|\beta|^2$$
$$= 2(|\alpha|^2 + |\beta|^2)$$

故

$$|\alpha+\beta|^2 + |\alpha-\beta|^2 = 2(|\alpha|^2 + |\beta|^2)$$

2）因为

$$\frac{1}{4}|\alpha+\beta|^2 - \frac{1}{4}|\alpha-\beta|^2 = \frac{1}{4}((\alpha+\beta),\ (\alpha+\beta))$$

$$-\frac{1}{4}((\alpha-\beta),\ (\alpha-\beta))$$

$$= \frac{1}{4}((\alpha\cdot\alpha) + 2(\alpha\cdot\beta) + (\beta\cdot\beta))$$

$$-\frac{1}{4}((\alpha\cdot\alpha) - 2(\alpha\cdot\beta) + (\beta\cdot\beta))$$

$$= \frac{1}{4}\cdot 4(\alpha\cdot\beta)$$

$$= (\alpha\cdot\beta).$$

3．因为

$$|\alpha+\beta|^2 = (\alpha+\beta,\ \alpha+\beta)$$
$$= (\alpha\cdot\alpha) + 2(\alpha\cdot\beta) + (\beta\cdot\beta)$$
$$= |\alpha|^2 + |\beta|^2$$

所以

$$|\alpha+\beta|^2 = |\alpha|^2 + |\beta|^2$$

4．设$\xi = (x_1, x_2, x_3, x_4)$与$\alpha, \beta, \gamma$正交，于是有方程组

$$x_1 + x_2 - x_3 + x_4 = 0$$
$$x_1 - x_2 - x_3 + x_4 = 0$$
$$2x_1 + x_2 + x_3 + 3x_4 = 0$$

解之得

$$\zeta=(4,\ 0,\ 1,-3)$$

再将其标准化

$$\eta=\frac{\zeta}{|\zeta|}=\left(\frac{4}{\sqrt{26}},0,\frac{1}{\sqrt{26}},\ -\frac{3}{\sqrt{26}}\right)$$

η 即为所求。

5．解：1）必要性，因为

$$\alpha=k\beta$$

所以

$$\cos\varphi=\frac{(\alpha\cdot\beta)}{|\alpha|\ |\beta|}=\frac{k(\beta\cdot\beta)}{|k|\ |\beta|\ |\beta|}=1$$

所以

$$\varphi=0$$

充分性，因为

$$\varphi=\cos^{-1}\frac{(\alpha\cdot\beta)}{|\alpha||\beta|}=0$$

所以

$$\frac{(\alpha\cdot\beta)}{|\alpha|\ |\beta|}=1$$

即

$$(\alpha\cdot\beta)=|\alpha||\beta|$$

从而α,β 线性相关，则有

$$\alpha=k\beta$$

于是有

$$\frac{(\alpha\cdot\beta)}{|\alpha||\beta|}=\frac{k(\beta\cdot\beta)}{|k|\ |\beta|\ |\beta|}=1$$

即

$$\frac{k}{|k|}=1$$

由此可知$k>0$。

2）必要性，因为

$$\alpha=k\beta$$

所以

$$\cos\varphi=\frac{(\alpha\cdot\beta)}{|\alpha|\ |\beta|}=\frac{k(\beta\cdot\beta)}{|k|\ |\beta|\ |\beta|}=-1$$

故

$$\varphi=\pi$$

充分性，因为

$$\varphi=\cos^{-1}\frac{(\alpha\cdot\beta)}{|\alpha|\ |\beta|}=\pi$$

所以

$$\frac{(\alpha\cdot\beta)}{|\alpha|\ |\beta|}=-1$$

即

$$(\alpha\cdot\beta)=-|\alpha|\cdot|\beta|$$

从而α,β 线性相关，即

$$\alpha=k\beta$$

于是

$$\frac{(\alpha,\beta)}{|\alpha|\ |\beta|}=\frac{k(\beta\cdot\beta)}{|k|\ |\beta|\ |\beta|}=\frac{k}{|k|}=-1$$

由此可知$k<0$。

6. 解：因为

$$\cos\varphi_i=\frac{(\alpha,\varepsilon_i)}{|\alpha|\ |\varepsilon_i|}=\frac{1}{\sqrt{n}\cdot 1}=\frac{1}{\sqrt{n}}$$

所以

$$\varphi_i=\cos^{-1}\frac{1}{\sqrt{n}}\qquad i=1,2,\cdots,n.$$

7. 证：设

$$\alpha=(a_1,a_2,\cdots,a_n),\quad \beta=(b_1,b_2,\cdots,b_n)$$

显然$\alpha,\beta\in D^{(n)}$，并且令当

$$当a_i=0时，\quad b_i=1$$

$$当a_i\neq 0时，\quad b_i=\frac{a_i}{|a_i|}$$

由$D^{(n)}$是欧氏空间，则有

$$(\alpha,\beta)^2\leqslant(\alpha,\alpha)\cdot(\beta,\beta)$$

$$(\alpha,\beta)=|a_1|+|a_2|+\cdots+|a_n|$$

$$(\alpha,\alpha)=a_1^2+a_2^2+\cdots+a_n^2$$

$$(\beta,\beta)=1^2+1^2+\cdots+1^2=n$$

即

$$\left(\sum_{i=1}^{n}|a_i|\right)^2\leqslant n(a_1^2+a_2^2+\cdots+a_n^2)$$

因此

$$\sum_{i=1}^{n}|a_i|\leqslant\sqrt{n(a_1^2+a_2^2+\cdots+a_n^2)}$$

8．证：1）因为

$$V=W\dotplus W^{\perp}=W^{\perp}\dotplus(W^{\perp})^{\perp}$$

对任意$\alpha\in W$，$\beta\in W^{\perp}$都有

$$(\alpha,\beta)=0$$

即

$$(\alpha,W^{\perp})=0$$

故$\alpha\in(W^{\perp})^{\perp}$，即$W\subseteq(W^{\perp})^{\perp}$，所以

$$(W^{\perp})^{\perp}=W.$$

2）对任意的$\alpha\in W_2^{\perp}$，$(\alpha,W_2)=0$，因$W_1\subseteq W_2$所以

$$(\alpha,W_1)=0$$

故$\alpha\in W_1^{\perp}$，因此有

$$W_2^{\perp}\subseteq W_1^{\perp}.$$

9. 证：因$\alpha_1, \alpha_2, \cdots, \alpha_n$是$V$的基底，且$\gamma \in V$，则有

$$\gamma = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$$

$$(\gamma, \gamma) = (k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n, \ \gamma)$$

$$= k_1(\alpha_1, \gamma) + k_2(\alpha_2, \gamma) + \cdots + k_n(\alpha_n, \gamma)$$

$$= 0$$

所以

$$\gamma = 0。$$

10. 证：因为

$$(\alpha_1, \beta) = (\alpha_2, \beta)$$

于是有

$$(\alpha_1 - \alpha_2, \beta) = 0$$

由于β 的任意性，则有

$$\alpha_1 - \alpha_2 = 0$$

故

$$\alpha_1 = \alpha_2。$$

练　习　二

1. 解：1）因为

$$(\alpha_1, \alpha_1) = 1, \quad (\alpha_1, \alpha_2) = 1, \quad (\alpha_1, \alpha_3) = 1$$

$$(\alpha_2, \alpha_1) = 1, \quad (\alpha_2, \alpha_2) = 2, \quad (\alpha_2, \alpha_3) = 2$$

$$(\alpha_3, \alpha_1) = 1, \quad (\alpha_3, \alpha_2) = 2, \quad (\alpha_3, \alpha_3) = 3$$

故$\alpha_1, \alpha_2, \alpha_3$的度量矩阵为

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

2）因为有可逆矩阵为

$$P=\begin{pmatrix}1 & -1 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}$$

使得

$$P'AP=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

故有

$$(\varepsilon_1,\varepsilon_2,\varepsilon_3)=(\alpha_1,\alpha_2,\alpha_3)P$$

于是得

$$\varepsilon_1=\alpha_1=(1,0,0)$$

$$\varepsilon_2=\alpha_2-\alpha_1=(0,1,0)$$

$$\varepsilon_3=\alpha_3-\alpha_2=(0,0,1)$$

$\varepsilon_1,\varepsilon_2,\varepsilon_3$是标准正交基底。

下面用标准正交化方法求标准正交基底，令

$$\varepsilon_1=\alpha_1=(1,0,0)$$

$$\varepsilon_2=\alpha_2-\frac{(\varepsilon_2,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1=(1,1,0)-(1,0,0)$$

$$=(0,1,0)$$

$$\varepsilon_3=\alpha_3-\frac{(\varepsilon_3,\varepsilon_1)}{(\varepsilon_1,\varepsilon_1)}\varepsilon_1-\frac{(\varepsilon_3,\varepsilon_2)}{(\varepsilon_2,\varepsilon_2)}\varepsilon_2$$

$$=(1,1,1)-(1,0,0)-(0,1,0)$$

$$=(0,0,1)$$

因$\varepsilon_1,\varepsilon_2,\varepsilon_3$ 是单位向量，所以$\varepsilon_1,\varepsilon_2,\varepsilon_3$是标准正交基底。

2．解：因为

$$\varepsilon_1=2\alpha_1+\alpha_2$$

$$\varepsilon_2=\alpha_1+\alpha_2$$

$$\varepsilon_3=\alpha_1+\alpha_3+\alpha_4$$

$$\varepsilon_4=-3\alpha_1-\alpha_2-\alpha_3+2\alpha_4$$

故由$\{\alpha_i\}$到$\{\varepsilon_i\}$的过渡矩阵

$$P=\begin{pmatrix}2 & 1 & 1 & -3\\ 1 & 1 & 0 & -1\\ 0 & 0 & 1 & -1\\ 0 & 0 & 1 & 2\end{pmatrix}$$

从而基底$\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$的度量矩阵为$A=P'BP=\begin{pmatrix}2 & 1 & 2 & -2\\ 1 & 2 & -1 & 0\\ 2 & -1 & 42 & 7\\ -1 & 0 & 7 & 3\end{pmatrix}$.

3．证：必要性

$$\begin{aligned}(\alpha, \varepsilon_i) &= (a_1\varepsilon_1+a_2\varepsilon_2+\cdots+a_n\varepsilon_n, \varepsilon_i)\\ &= a_1(\varepsilon_1, \varepsilon_i)+a_2(\varepsilon_2, \varepsilon_i)+\cdots+a_n(\varepsilon_n, \varepsilon_i)\\ &= a_i\end{aligned}$$

充分性，显然ε_i在$\{\varepsilon_i\}$上的坐标为

$$(0,\cdots,0,\underset{(i)}{1},0,\cdots,0)$$

由$(\alpha, \varepsilon_i)=a_i$有

$$(\varepsilon_i, \varepsilon_j)=\begin{cases}0, & i\neq j\\ 1, & i=j\end{cases}$$

故$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$是$V$的标准正交基底。

4．解：由齐次线性方程组

$$2x_1+x_2-x_3+x_4-3x_5=0$$

$$x_1+x_2-x_3+x_5=0$$

得解空间的基础解系为

$$\alpha_1=(1, 0, 0, -5, -1)$$

$$\alpha_2=(0, 1, 0, -4, -1)$$

$$\alpha_3=(0, 0, 1, 4, 1)$$

将其正交化

$$\beta_1=(1, 0, 0, -5, -1)$$

$$\beta_2=\alpha_2-\frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1$$

$$=\frac{1}{9}(-7,9,0,-1,-2)$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2$$

$$=\frac{1}{15}(7,6,15,1,2)$$

再标准化

$$\eta_1=\frac{1}{3\sqrt{3}}(1,0,0,-5,-1)$$

$$\eta_2=\frac{1}{3\sqrt{15}}(-7,9,0,-1,-2)$$

$$\eta_3=\frac{1}{3\sqrt{35}}(7,6,15,1,2)$$

是解空间的标准正交基底。

5．证：将α_1，α_2，…，α_m 扩为V的标准正交基底：$\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n$。令

$$\alpha=x_1\alpha_1+x_2\alpha_2+\cdots+x_m\alpha_m+x_{m+1}\alpha_{m+1}+\cdots+x_n\alpha_n$$

由第3题有

$$\sum_{i=1}^{m}(\alpha,\alpha_i)^2=x^2{}_1+x^2{}_2+\cdots+x^2{}_m$$

$$|\alpha|=(\alpha,\alpha)=x^2{}_1+x_2^2+\cdots+x^2{}_m+x^2{}_{m+1}+\cdots+x_n^2$$

故

$$\sum_{i=1}^{m}(\alpha,\alpha_i)^2\leqslant|\alpha|^2.$$

6．因为由$\{\varepsilon_i\}$到$\{\alpha_i\}$的过渡矩阵为

$$T=\begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{pmatrix}$$

T为正交阵，因此α_1，α_2，α_3是标准正交基底。

7. 解：容易证明α_1,α_2，α_3是线性无关的，故是V_1的基底，将其正交化得

$$\beta_1=\varepsilon_1+\varepsilon_5$$

$$\beta_2=\frac{1}{2}\varepsilon_1-\varepsilon_2+\varepsilon_4-\frac{1}{2}\varepsilon_5$$

$$\beta_3=\varepsilon_1+\varepsilon_2+\varepsilon_3-\varepsilon_5$$

再标准化得

$$\eta_1=\frac{\sqrt{2}}{2}(\varepsilon_1+\varepsilon_5)$$

$$\eta_2=\frac{\sqrt{10}}{5}\ (\varepsilon_1-2\varepsilon_2+2\varepsilon_4-\varepsilon_5)$$

$$\eta_3=\frac{1}{2}(\varepsilon_1+\varepsilon_2+\varepsilon_3-\varepsilon_5)$$

故η_1,η_2,η_3 是V_1的的标准正交基底。

练　习　三

1. 证：设σ_1，σ_2是正交变换，$\{\alpha_i\}$是标准正交基底，σ_1,σ_2在$\{\alpha_i\}$上的表示矩阵为T_1,T_2，那么$\sigma_1\sigma_2$在$\{\sigma_i\}$上的表示矩阵为T_1T_2，因T_1，T_2是正交阵，而正交阵的乘积仍为正交阵，故 T_1T_2 是正交阵，因此$\sigma_1\cdot\sigma_2$是正交变换。

又因σ^{-1}在$\{\alpha_i\}$上的阵为 T^{-1}，且正交阵的逆阵仍为正交阵，故σ^{-1}仍是正交变换。

2．证：1）对任意α,β

$$\sigma(k\alpha+l\beta)=k\alpha+l\beta-2(\eta,k\alpha+l\beta)\cdot\eta$$
$$=k\sigma(\alpha)+l\sigma(\beta)$$

故σ是线性变换．

$$(\sigma(\alpha),\sigma(\beta))=(\alpha-2(\eta,\alpha)\eta,\beta-2(\eta,\beta)\eta)$$
$$=(\alpha,\beta)-2(\eta,\beta)(\eta,\alpha)-2(\eta,\alpha)(\eta,\beta)$$
$$+4(\eta,\alpha)(\eta,\beta)(\eta,\eta)$$
$$=(\alpha,\beta)$$

故σ是正交变换．

2）由η是单位向量，将η扩为标准正交基底，

$$\eta,\ \varepsilon_2,\varepsilon_3,\cdots,\varepsilon_n$$

因为

$$\sigma(\eta)=\eta-2(\eta,\eta)\cdot\eta=-\eta$$
$$\sigma(\varepsilon_i)=\varepsilon_i-2(\eta,\varepsilon_i)\cdot\eta=\varepsilon_i\quad i=2,3,\cdots,\ n.$$

所以σ在η，$\varepsilon_2,\varepsilon_3,\cdots,\varepsilon_n$上的表示矩阵为

$$A=\begin{pmatrix}-1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{pmatrix}$$

因此

$$|A|=-1$$

3）σ的特征根有n个，现有$n-1$个1，另一个也一定是实数，设为λ_0，于是有

$$\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}\lambda_0 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1\end{pmatrix}$$

于是有

$$(\sigma(\varepsilon_1),\sigma(\varepsilon_1))=\lambda_0^2(\varepsilon_1,\varepsilon_1)=(\varepsilon_1,\varepsilon_1)$$

故

$$\lambda_0=\pm1$$

但因V_1是$n-1$维的，因此

$$\lambda_0 = -1$$

于是

$$\sigma(\varepsilon_1) = -\varepsilon_1,$$

$$\sigma(\varepsilon_i) = \varepsilon_i \qquad i = 2, 3, \cdots, n$$

进而有

$$(\varepsilon_1, \varepsilon_i) = 0 \qquad i = 2, 3, \cdots, n$$

现在令，

$$\eta = \frac{\varepsilon_1}{|\varepsilon_1|}$$

那么

$$(\eta, \varepsilon_i) = 0, \quad i = 2, 3, \cdots, n$$

$$\sigma(\eta) = \sigma\left(\frac{\varepsilon_1}{|\varepsilon_1|}\right) = \frac{1}{|\varepsilon_1|}\sigma(\varepsilon_1) = \frac{1}{|\varepsilon_1|}(-\varepsilon_1) = -\eta$$

下面证明对任一$\alpha \in V$ 都有

$$\sigma(\alpha) = \alpha - 2(\alpha, \eta)\eta$$

任取$\alpha \in V$，则有

$$\alpha = k_1\eta + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n$$

$$\begin{aligned}\sigma(\alpha) &= \sigma(k_1\eta + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n)\\ &= k_1\sigma(\eta) + k_2\sigma(\varepsilon_2) + \cdots + k_n\sigma(\varepsilon_n)\\ &= -k_1\eta + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n\end{aligned}$$

$$\begin{aligned}2(\alpha, \eta)\eta &= 2(k_1\eta + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n, \ \eta)\eta\\ &= 2k_1\eta\end{aligned}$$

$$\alpha - 2(\alpha, \eta)\eta = -k_1\eta + k_2\varepsilon_2 + \cdots + k_n\varepsilon_n$$

因此

$$\sigma(\alpha) = \alpha - 2(\alpha, \eta)\cdot\eta.$$

3．证明：将$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m$扩为$V$的标准正交基底 ε_1，ε_2，$\cdots$，$\varepsilon_m, \varepsilon_{m+1}, \cdots$，$\varepsilon_n$. 同时将$\alpha_1, \alpha_2, \cdots, \alpha_m$也扩为$V$ 的标准正交基底，即$\alpha_1, \alpha_2, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$. 则由线性变换一章的习题第1题有，存在

线性变换σ 使得，

$$\sigma(\varepsilon_i)=\alpha_i \qquad i=1,2,\cdots,m,m+1,\cdots,n,$$

又因σ将标准正交基底$\{\varepsilon_i\}$变为标准正交基底$\{\alpha_i\}$，故σ是正交变换．

4．证：令

$$e_1=\frac{\alpha}{|\alpha|}$$

并将e_1扩为标准正交基底，

$$e_1,e_2,\cdots,e_n$$

再令

$$f_1=\frac{\beta}{|\beta|}$$

并将f_1扩为标准正交基底，

$$f_1,f_2,\cdots,f_n$$

由上题知，存在正交变换σ，使得

$$\sigma(e_i)=f_i.$$

则

$$\sigma(e_1)=\sigma\left(\frac{\alpha}{|\alpha|}\right)=\frac{1}{|\alpha|}\sigma(\alpha)=f_1=\frac{\beta}{|\beta|}$$

又因为α,β 的长相同，即

$$|\alpha|=|\beta|$$

因此有

$$\sigma(\alpha)=\beta$$

5．证：因为

$$\begin{aligned}&(\sigma(\alpha+\beta)-\sigma(\alpha)-\sigma(\beta),\ \sigma(\alpha+\beta)-\sigma(\alpha)-\alpha(\beta))\\&=(\sigma(\alpha+\beta),\sigma(\alpha+\beta))-2(\sigma(\alpha+\beta),\sigma(\alpha))-2(\sigma(\alpha+\beta),\\&\quad\sigma(\beta))+(\sigma(\alpha),\sigma(\alpha))+(\sigma(\beta),\sigma(\beta))+2(\sigma(\alpha),\sigma(\beta))\\&=((\alpha+\beta),(\alpha+\beta))-2((\alpha+\beta),\alpha)-2(\alpha+\beta,\beta)+(\alpha,\alpha)\\&\quad+(\beta,\beta)+2(\alpha,\beta)\\&=(\alpha+\beta,\ \alpha+\beta)-2(\alpha+\beta,\alpha+\beta)+(\alpha+\beta,\alpha+\beta)\\&=0\end{aligned}$$

所以

$$\sigma(\alpha+\beta)-\sigma(\alpha)-\sigma(\beta)=0$$

即

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta)$$

又

$$\begin{aligned}&(\sigma(k\alpha)-k\sigma(\alpha),\sigma(k\alpha)-k\sigma(\alpha))\\&=(\sigma(k\alpha),\sigma(k\alpha))-2(k\sigma(\alpha),\sigma(k\alpha))+(k\sigma(\alpha),k\sigma(\alpha))\\&=(k\alpha,k\alpha)-2k(\alpha,k\alpha)+k^2(\alpha,\alpha)\\&=k^2(\alpha,\alpha)-2k^2(\alpha,\alpha)+k^2(\alpha,\alpha)\\&=0\end{aligned}$$

所以

$$\sigma(k\alpha)-k\sigma(\alpha)=0$$

即

$$\sigma(k\alpha)=k\sigma(\alpha)$$

因此σ 是线性变换。又因为

$$(\sigma(\alpha),\sigma(\beta))=(\alpha,\beta)$$

因此σ 是正交变换。

6. 证：取$e_1,e_2,\cdots,e_m$为W 的标准正交基底，并扩为V的标准正交基底

$$e_1,e_2,\cdots,\ e_m,e_{m+1},\ e_{m+2},\cdots,\ e_n$$

显然有

$$W^{\perp}=L(e_{m+1},\ e_{m+2},\cdots,\ e_n)$$

因正交变换将标准正交基底变为标准正交基底，故

$$\sigma(e_1),\sigma(e_2),\cdots,\ \sigma(e_m),\sigma(e_{m+1}),\cdots,\sigma(e_n)$$

仍是V的标准正交基底，而

$$\sigma(e_1),\sigma(e_2),\cdots,\ \sigma(e_m)\in W$$

$$\sigma(e_{m+1}),\ \sigma(e_{m+2}),\cdots,\ \sigma(e_n)\in W^{\perp}$$

且是他们的基底，因此$W^{\perp}$是σ 的不变子空间。

7. 证：因为对任意$\alpha,\beta\in V$

$$((\sigma_1\sigma_2+\sigma_2\sigma_1)(\alpha),\beta)=(\sigma_1\sigma_2(\alpha),\beta)+(\sigma_2\sigma_1(\sigma),\beta)$$
$$=(\sigma_2(\alpha),\sigma_1(\beta))+(\sigma_1(\alpha),\sigma_2(\beta))$$
$$=(\alpha,\sigma_2\sigma_1(\beta))+(\alpha,\sigma_1\sigma_2(\beta))$$
$$=(\alpha,\sigma_2\sigma_1(\beta)+\sigma_1\sigma_2(\beta))$$
$$=(\alpha,(\sigma_2\sigma_1+\sigma_1\sigma_2)(\beta))$$

故$\sigma_2\sigma_1+\sigma_1\sigma_2$是对称变换。

8. 证：设$\alpha\in W$，$\beta\in W^{\perp}$，因$\sigma(\alpha)\in W$，故有

$$(\sigma(\alpha),\beta)=(\alpha,\sigma(\beta))=0$$

故

$$\sigma(\beta)\in W^{\perp}$$

因此$W^{\perp}$是σ的不变子空间。

习 题 十三

1. 解：令$\alpha_1=1,\alpha_2=x$，$\alpha_3=x^2$，$\alpha_4=x^3$

将其正交化，设

$$\beta_1=\alpha_1=1$$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}$$

因为

$$(\alpha_2,\beta_1)=\int_{-1}^{1}x\cdot 1d(x)=0$$

所以

$$\beta_2=x$$

$$\beta_3=\alpha_3-\frac{(\alpha_3,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_3,\beta_2)}{(\beta_2,\beta_2)}\beta_2$$

因为

$$(\alpha_3, \beta_1) = \int_{-1}^{1} x^2 dx = \frac{2}{3}$$

$$(\alpha_3, \beta_2) = \int_{-1}^{1} x^2 x dx = 0$$

$$(\beta_1, \beta_1) = \int_{-1}^{1} 1 \ dx = 2$$

所以

$$\beta_3 = x^2 - \frac{1}{3}$$

$$\beta_4 = \alpha_4 - \frac{(\alpha_4, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \frac{(\alpha_4, \beta_2)}{(\beta_2, \beta_2)}\beta_2 - \frac{(\alpha_4, \beta_3)}{(\beta_3, \beta_3)}\beta_3$$

因为

$$(\alpha_4, \beta_1) = \int_{-1}^{1} x^3 dx = 0$$

$$(\alpha_4 \beta_2) = \int_{-1}^{1} x^4 dx = \frac{2}{5}$$

$$(\beta_2, \beta_2) = \int_{1}^{-1} x^2 dx = \frac{2}{3}$$

$$(\alpha_4, \beta_3) = \int_{-1}^{1} x^5 dx = 0$$

故

$$\beta_4 = x^3 - \frac{3}{5}x$$

再将其标准化

$$\eta_1 = \frac{1}{|\beta_1|}\beta_1 = \frac{\sqrt{2}}{2}$$

$$\eta_2 = \frac{1}{|\beta_2|}\beta_2 = \frac{\sqrt{6}}{2}x$$

$$\eta_3 = \frac{1}{|\beta_3|}\beta_3 = \frac{\sqrt{10}}{4}(3x^2 - 1)$$

$$\eta_4 = \frac{1}{|\beta_4|}\beta_4 = \frac{\sqrt{14}}{4}(15x^3 - 3x)$$

由此而得$\eta_1, \eta_2, \eta_3, \eta_4$是$D_4[x]$中一组标准正交基底。

2. 证：由于

$\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$与$\frac{2(\alpha,\beta)}{(\beta,\beta)}$都是不大于0的整数，故有

$$(\alpha,\beta)\leqslant 0$$

又由于

$$\frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)}=\frac{4(\alpha,\beta)^2}{|\alpha|^2|\beta|^2}\geqslant 0 \text{ 的正整数。}$$

由柯西不等式有

$$\frac{|(\alpha,\beta)|}{|\alpha|\ |\beta|}\leqslant 1$$

所以

$$\frac{(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)}\leqslant 1$$

故有

$$0\leqslant\frac{4(\alpha,\beta)^2}{(\alpha,\alpha)(\beta,\beta)}\leqslant 4 \text{ 且是整数}$$

设α,β之间夹角为θ，则有

$$0\leqslant 4\cos^2\theta\leqslant 4$$

因α,β线性无关，故$\cos\theta\neq\pm 1$，且$4\cos^2\theta$是整数，所以有

$$4\cos^2\theta=0,$$

$$4\cos^2\theta=1$$

$$4\cos^2\theta=2$$

$$4\cos^2\theta=3$$

由$(\alpha,\beta)\leqslant 0$，则有

$$\cos\theta=0 \qquad \theta=\frac{\pi}{2}$$

$$\cos\theta=-\frac{1}{2} \qquad \theta=\frac{2}{3}\pi$$

$$\cos\theta = -\frac{\sqrt{2}}{2} \qquad \theta = \frac{3}{4}\pi$$

$$\cos\theta = -\frac{\sqrt{3}}{2} \qquad \theta = \frac{5}{6}\pi$$

因此α,β 的夹角只能是$\frac{\pi}{2}$，$\frac{2}{3}\pi,\frac{3}{4}\pi,\frac{5}{6}\pi$。

3．证：1）设$\alpha\in(W_1+W_2)^{\perp}$即

$$(\alpha, W_1+W_2)=0 \tag{1}$$

对任意$\beta\in W_1$，

$$\beta=\beta+0 \ \in W_1+W_2$$

则由（I）得

$$(\alpha,\beta)=0$$

即

$$\alpha\in W_1^{\perp}$$

同理可得

$$\alpha\in W_2^{\perp}$$

故

$$\alpha\in W_1^{\perp}\cap W_2^{\perp} \tag{2}$$

又任取$\alpha\in W_1^{\perp}\cap W_2^{\perp}$，则有

$$(\alpha, W_1)=0,\ (\alpha, W_2)=0$$

任取$\beta\in W_1+W_2$，则

$$\beta=\beta_1+\beta_2,\quad \beta_i\in W_i,\quad i=1,2.$$

$$(\alpha,\beta)=(\alpha,\ \beta_1+\beta_2)=(\alpha,\beta_1)+(\alpha,\beta_2)=0$$

故

$$\alpha\in(W_1+W_2)^{\perp}$$

因此

$$(W_1+W_2)^{\perp}\supseteq W_1^{\perp}\cap W_2^{\perp}$$

故

$$(W_1+W_2)^{\perp}=W_1^{\perp}\cap W_2^{\perp} \tag{3}$$

2）将（3）中的W_1，W_2换成$W_1^{\perp}$，与$W_2^{\perp}$则有

$$(W_1^{\perp}+W_2^{\perp})^{\perp}=(W_1^{\perp})^{\perp}\cap(W_2^{\perp})^{\perp}=W_1\cap W_2$$

故有

$$W_1^{\perp}+W_2^{\perp}=(W_1\cap W_2)^{\perp}$$

4．证：考查等式

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n=0 \tag{1}$$

由（1）分别与$\alpha_1,\alpha_2,\cdots,\alpha_n$取数积则有

$$\begin{cases} k_1(\alpha_1,\alpha_1)+k_2(\alpha_1,\alpha_2)+\cdots+k_n(\alpha_1,\alpha_n)=0 \\ k_1(\alpha_2,\alpha_1)+k_2(\alpha_2,\alpha_2)+\cdots+k_n(\alpha_2,\alpha_n)=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ k_1(\alpha_n,\alpha_1)+k_2(\alpha_n,\alpha_2)+\cdots+k_n(\alpha_n,\alpha_n)=0 \end{cases} \tag{2}$$

由上式知 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关当且仅当（2）只有零解当且仅当

$G(\alpha_1,\alpha_2,\cdots,\alpha_n)\neq 0$。

5．证：因为$\varepsilon_1,\varepsilon_2\cdots,\varepsilon_n$是标准正交基底，所以

$$(\alpha_i,\alpha_j)=a_{i1}a_{j1}+\cdots+a_{in}a_{jn}=\sum_{k=1}^{n}a_{ik}\,a_{jk}$$

$$G(\alpha_1,\alpha_2,\cdots,\alpha_n)=\begin{vmatrix} \sum_{k=1}^{n}a_{1k}a_{1k} & \sum_{k=1}^{n}a_{1k}a_{2k}\cdots & \sum_{k=1}^{n}a_{1k}a_{nk} \\ \sum_{k=1}^{n}a_{2k}a_{1k} & \sum_{k=1}^{n}a_{2k}a_{2k}\cdots & \sum_{k=1}^{n}a_{2k}a_{nk} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \sum_{k=1}^{n}a_{nk}a_{1k} & \sum_{k=1}^{n}a_{nk}a_{2k}\cdots & \sum_{k=1}^{n}a_{nk}a_{nk} \end{vmatrix}$$

$$=\begin{vmatrix} a_{11} & a_{12}\cdots a_{1n} \\ a_{21} & a_{22}\cdots a_{2n} \\ \cdots\cdots & \cdots\cdots \\ a_{n1} & a_{n2}\cdots a_{nn} \end{vmatrix}\cdot\begin{vmatrix} a_{11} & a_{21}\cdots a_{n1} \\ a_{12} & a_{22}\cdots a_{n2} \\ \cdots\cdots & \cdots\cdots \\ a_{1n} & a_{2n}\cdots a_{nn} \end{vmatrix}$$

$$=|a_{ij}|^2$$

因此

$$G(\alpha_1, \alpha_2, \cdots, \alpha_n) = |a_{ij}|^2$$

6．证：将$\beta_1, \beta_2, \cdots, \beta_n$标准化而得到标准正交基底

$$\varepsilon_i = \frac{\beta_i}{|\beta_i|} \qquad i = 1, 2, \cdots, n$$

即

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} |\beta_1| & & & \\ & |\beta_2| & & \\ & & \ddots & \\ & & & |\beta_n| \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

由上题知

$$\begin{aligned} G(\beta_1, \beta_2, \cdots, \beta_n) &= (|\beta_1| \ |\beta_2| \cdots |\beta_n|)^2 \\ &= (\beta_1, \beta_1)(\beta_2, \beta_2)\cdots(\beta_n, \beta_n) \end{aligned}$$

因β_1，$\beta_2, \cdots$，β_n是由$\alpha_1, \alpha_2, \cdots$，$\alpha_n$正交化得到的，故

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = M \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

M是下三角阵，且$|M| = 1$，于是有

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = M^{-1} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

故M^{-1}也是下三角阵，且$|M^{-1}| = 1$，进而有

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = M^{-1} \begin{pmatrix} |\beta_1| & & & \\ & |\beta_2| & & \\ & & \ddots & \\ & & & |\beta_n| \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

故

$$\begin{aligned} G(\alpha_1, \alpha_2, \cdots, \alpha_n) &= \left| M^{-1} \begin{pmatrix} |\beta_1| & & & \\ & |\beta_2| & & \\ & & \ddots & \\ & & & |\beta_n| \end{pmatrix} \right|^2 \\ &= (\beta_1, \beta_1)(\beta_2, \beta_2)\cdots(\beta_n, \beta_n) \end{aligned}$$

因此有

$$G(\alpha_1,\alpha_2,\cdots,\alpha_n)=G(\beta_1,\beta_2,\cdots,\beta_n)=(\beta_1,\beta_1)(\beta_2,\beta_2)\cdots(\beta_n,\beta_n) .$$

7．证：若$S_1\cap S_2$不空，设$\alpha\in S_1\cap S_2$，则有

$$(\alpha,\alpha)=0$$

故有

$$\alpha=0.$$

8．证：对任意$\xi\in W_1$，$\eta\in W_2$，于是有

$$\xi=\sum_{i=1}^{m}x_i\alpha_i$$

$$\eta=\sum_{j=1}^{s}y_j\beta_j$$

$$(\xi,\eta)=\left(\sum_{i=1}^{m}x_i\alpha_i,\ \sum_{j=1}^{s}y_j\beta_j\right)=\sum_{i=1}^{m}\sum_{j=1}^{s}x_iy_j\ (\alpha_i,\beta_j)$$

$$=0$$

故W_1，W_2正交。

9．证：设

$$\alpha=\alpha_1+\alpha_2$$

其中$\alpha_1\in W$，$\alpha_2\in W^{\perp}$，令$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$是$W$的标准正交基底，则有

$$\alpha_1=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_m\varepsilon_m$$

于是有

$$\alpha=k_1\varepsilon_1+k_2\varepsilon_2+\cdots+k_m\varepsilon_m+\alpha_2$$

而且

$$(\alpha,\varepsilon_i)=k_i,\quad i=1,2,\cdots,m$$

故

$$\alpha_1=(\alpha,\varepsilon_1)\varepsilon_1+(\alpha,\varepsilon_2)\varepsilon_2+\cdots+(\alpha,\varepsilon_m)\varepsilon_m.$$

10．证明：必要性：对任意$\xi\in W$，则有

$$\alpha-\xi=\alpha-\beta+\beta-\xi$$

$\beta-\xi\in W$，而$\alpha-\beta\in W^{\perp}$，所以

$$(\alpha-\beta,\beta-\xi)=0$$

$$|\alpha-\xi|^2=(\alpha-\xi,\alpha-\xi)$$
$$=(\alpha-\beta+\beta-\xi,\ \alpha-\beta+\beta-\xi)$$
$$=(\alpha-\beta,\alpha-\beta)+(\beta-\xi,\beta-\xi)$$
$$=|\alpha-\beta|^2+|\beta-\xi|^2$$

又因为

$$|\beta-\xi|^2\geqslant 0$$

故有

$$|\alpha-\xi|^2\geqslant|\alpha-\beta|^2$$

因此

$$|\alpha-\xi|\geqslant|\alpha-\beta|$$

充分性：设β_1是α在W上的内射影，由必要性证明知

$$|\alpha-\beta_1|\leqslant|\alpha-\beta|$$

又由已知条件

$$|\alpha-\beta|\leqslant|\alpha-\beta_1|$$

故有

$$|\alpha-\beta|=|\alpha-\beta_1|$$

即

$$(\alpha-\beta,\alpha-\beta)=(\alpha-\beta_1,\ \alpha-\beta_1)$$

令

$$\alpha=\beta_1+\gamma$$

于是

$$(\alpha-\beta,\alpha-\beta)=(\beta_1-\beta+\gamma,\ \beta_1-\beta+\gamma)$$
$$=(\beta_1-\beta,\beta_1-\beta)+(\gamma,\gamma)$$
$$=(\beta_1-\beta,\beta_1-\beta)+(\alpha-\beta_1,\alpha-\beta_1)$$

故有

$$(\beta_1-\beta,\beta_1-\beta)=0$$

因此有

$$\beta_1=\beta$$

即β 是α 在W上的内射影

11. 证：因A是正交变换σ 在标准正交基底上的表示矩阵，故A是正交阵。又因为

$$|(-1)I-A| = |(-1)AA'-A| = |A||-A'-I|$$
$$=|A'||-I-A'|$$
$$=|A||-I-A|$$

又因为$|A|=-1$,故有

$$|(-1)I-A| = -|(-1)I-A|$$

所以有

$$|(-1)I-A| = 0$$

即(-1)是σ 的特征根。

12. 证：因α_1，α_2，…，α_n是标准正交基底，τ 是正交变换，则$\tau(\alpha_1),\tau(\alpha_2),\cdots$，$\tau(\alpha_n)$也是标准正交基底，于是有

$$(\tau(\alpha_1),\tau(\alpha_2),\cdots,\tau(\alpha_n)) = (\beta_1,\beta_2,\cdots,\beta_n)T$$

显然T是正交阵，令

$$T=(a_{ij})$$

由$\tau(\alpha_1)=\beta_1$,可得

$$T=\begin{pmatrix} 1 & a_{12}\cdots a_{1n} \\ 0 & \\ \vdots & T_{n-1} \\ 0 & \end{pmatrix}$$

由T是正交阵可有

$$1^2+a^2_{12}+\cdots+a^2_{1n}=1$$

由此而得$a_{1k}=0$，$(k=2,3,\cdots,n)$。

于是

$$T=\begin{pmatrix} 1 & 0\cdots 0 \\ 0 & \\ \vdots & T_{n-1} \\ 0 & \end{pmatrix}$$

即

$$\tau(\alpha_i)=a_{2i}\beta_2+a_{3i}\beta_3+\cdots+a_{ni}\beta_n \qquad i=2\cdots,n$$

所以

$$L(\tau(\alpha_2),\tau(\alpha_3),\cdots,\tau(\alpha_n))\subseteq L(\beta_2,\cdots,\beta_n)$$

又因为

$$\dim L(\tau(\alpha_2),\tau(\alpha_3),\cdots,\tau(\alpha_n))=\dim L(\beta_2,\cdots,\beta_n)$$

故

$$L(\tau(\alpha_2),\tau(\alpha_3),\cdots,\tau(\alpha_n))=L(\beta_2,\cdots,\beta_n).$$

13．证：1）必要性，设σ 在标准正交基底 $e_1,e_2,\cdots,e_n$ 上的表示矩阵为反对称阵，则

$$(\sigma(e_i),e_j)=(-e_i,\sigma(e_j))$$

对任意 α，$\beta\in V$，有

$$\alpha=c_1e_1+c_2e_2+\cdots+c_ne_n=\sum_{i=1}^{n}c_ie_i$$

$$\beta=b_1e_1+b_2e_2+\cdots+b_ne_n=\sum_{i=1}^{n}b_ie_i$$

$$(\sigma(\alpha),\beta)=\left(\sigma\sum_{i=1}^{n}c_ie_i,\sum_{i=1}^{n}b_ie_i\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}c_ib_j\ (\sigma(e_i),e_j)$$

同理可有

$$(\alpha,\sigma(\beta))=\sum_{i=1}^{n}\sum_{j=1}^{n}c_ib_j(e_i,\sigma(e_j))$$

故

$$(\sigma(\alpha),\beta)=-(\alpha,\sigma(\beta))$$

充分性，设$e_1,e_2,\cdots,e_n$ 是一个标准正交基底

$$\sigma(e_i)=k_{1i}e_1+k_{2i}e_2+\cdots+k_{ni}e_n\qquad i=1,2,\cdots,n$$

于是

$$(\sigma(e_i),e_j)=k_{ji}$$

$$(\sigma(e_j),e_i)=k_{ij}$$

由反对称性知

$$(\sigma(e_i), e_j) = -(e_i, \sigma(e_j))$$

$$k_{ji} = -k_{ij}$$

故σ 在标准正交基底$e_1, e_2, \cdots, e_n$ 上的表示矩阵为反对称阵，因此σ是反对称变换.

2）对任意$\alpha \in W_1^{\perp}$，去证明 $\sigma(\alpha) \in W_1^{\perp}$，

任取$\beta \in W_1$，有$\sigma(\beta) \in W_1$，则

$$(\alpha, \sigma(\beta)) = 0$$

由题设σ是反对称的，所以

$$(\sigma(\alpha), \beta) = -(\alpha, \sigma(\beta)) = 0$$

故$\sigma(\alpha) \in W_1^{\perp}$，$W_1^{\perp}$是$\sigma$ 的不变子空间.

14.证：1）由练习三的第2题知，镜面反射变为α

$$\sigma(\alpha) = \alpha - 2(\eta, \alpha)\eta \qquad \eta \text{ 为单位向量} \quad (1)$$

令

$$\beta = \alpha - 2(\eta, \alpha)\eta \qquad (2)$$

$$\alpha - \beta = 2(\eta, \alpha)\eta$$

因为$\alpha \neq \beta$，所以

$$(\eta, \alpha) \neq 0$$

于是

$$\eta = \frac{\alpha - \beta}{2(\eta, \alpha)} \qquad (3)$$

面

$$(\eta, \alpha) = \left(\frac{\alpha - \beta}{2(\eta, \alpha)}, \alpha\right)$$

$$= \frac{1}{2(\eta, \alpha)}[(\alpha, \alpha) - (\beta, \alpha)]$$

$$= \frac{1}{2(\eta, \alpha)}(1 - (\beta, \alpha))$$

$$(\alpha, \eta)^2 = \frac{1}{2}(1 - (\alpha, \beta))$$

$$(\alpha,\eta)=\frac{1}{2}\sqrt{2(1-(\alpha,\beta))} \qquad (4)$$

将（4）代入（3）得

$$\eta=\frac{\alpha-\beta}{\sqrt{2(1-(\alpha,\beta))}} \qquad (5)$$

且可验证

$$(\eta,\eta)=1$$

再将（5）代入镜面反射公式，且由（2）可知镜面反射

$$\sigma(\alpha)=\beta$$

2）设σ 是正交变换，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是一个标准正交基底，

$$\sigma(\varepsilon_i)=\eta_i \qquad i=1,2,\cdots,n.$$

则$\eta_1,\eta_2,\cdots,\eta_n$ 是标准正交基底．如果

$$\varepsilon_i=\eta_i \qquad i=1,2,\cdots,n$$

此时，我们定义

$$\sigma_1(\varepsilon_1)=-\varepsilon_1$$

$$\sigma_1(\varepsilon_i)=\varepsilon_i \qquad i=2,3,\cdots,n$$

可以验证σ_1是镜面反射，且有

$$\sigma=\sigma_1\sigma_1$$

下面若$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$与$\eta_1,\eta_2,\cdots,\eta_n$不尽相同，不妨设$\varepsilon_1\neq\eta_1$，由1）知，存在镜面反射$\sigma_1$

$$\sigma_1(\varepsilon_1)=\eta_1$$

$$\sigma_1(\varepsilon_i)=\xi_i \qquad i=2,3,\cdots,n$$

即

$$\sigma_1(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\eta_1,\xi_2\cdots,\xi_n)$$

这里若$\xi_i=\eta_i$，则$\sigma=\sigma_1$问题得证．否则，若$\xi_2\neq\eta_2$，又可定义

$$\sigma_2(\alpha)=\alpha-2(\eta,\alpha)\eta$$

其中

$$\eta=\frac{\xi_2-\eta_2}{\sqrt{2(\xi_2,\eta_2)}}$$

那么可以验证

$$\sigma_2(\xi_2)=\eta_2$$

$$\sigma_2(\eta_1)=\eta_1$$

则有

$$\sigma_2\sigma_1(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\eta_1,\eta_2,\xi_3{}',\cdots,\xi'_n)$$

仿此做下去，则有

$$\sigma_s\cdots\sigma_1(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\eta_1\eta_2\cdots\eta_n)$$

所以有

$$\sigma=\sigma_t\cdots\sigma_2\sigma_1$$

其中σ_i是镜面反射。

15．必要性：因σ 是正交变换，且

$$\sigma(\alpha_i)=\beta_i$$

故有

$$(\alpha_i,\alpha_j)=(\sigma(\alpha_i),\sigma(\alpha_j))=(\beta_i,\beta_j)$$

充分性：分四步进行证明

首先，不妨设$\alpha_1,\alpha_2,\cdots,\alpha_m$的极大无关组为$\alpha_1,\alpha_2,\cdots,\alpha_r$；去证 $\beta_1,\beta_2,\cdots,\beta_r$也是极大无关组。

事实上，由 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 是线性无关，据第4题有

$$G(\alpha_1,\alpha,\cdots,\alpha_r)=\begin{vmatrix}(\alpha_1,\alpha_1)&(\alpha_1,\alpha_2)&\cdots&(\alpha_1,\alpha_r)\\(\alpha_2,\alpha_1)&(\alpha_2,\alpha_2)&\cdots&(\alpha_2,\alpha_r)\\\cdots&\cdots&\cdots&\cdots\\(\alpha_r,\alpha_1)&(\alpha_r,\alpha_2)&\cdots&(\alpha_r,\alpha_r)\end{vmatrix}\neq0$$

又由

$$(\alpha_i,\alpha_j)=(\beta_i,\beta_j)$$

有$G(\beta_1,\beta_2,\cdots,\beta_r)\neq0$，因此$\beta_1,\beta_2,\cdots,\beta_r$线性无关。

又在$\beta_1,\beta_2,\cdots,\beta_m$中任取$\beta_s$，考虑 α_s，由$\alpha_1,\alpha_2,\cdots,\alpha_r$ 是极大无关组，则有

$$\alpha_s=\sum_{i=1}^{r}k_i\alpha_i$$

那么

$$\left(\beta_s-\sum_{i=1}^{r}k_i\beta_i,\beta_s-\sum_{i=1}^{r}k_i\beta_i\right)=(\beta_s,\beta_s)-2\left(\beta_s,\ \sum_{i=1}^{r}k_i\beta_i\right)$$

$$+\left(\sum_{i=1}^{r}k_i\beta_i,\sum_{i=1}^{r}k_i\beta_i\right)$$

$$=(\beta_s,\beta_s)-2\ \sum_{i=1}^{r}(k_i(\beta_s,\beta_i))+\sum_{j=1}^{r}\sum_{i=1}^{r}k_ik_j\ (\beta_i,\beta_j)$$

$$=(\alpha_s,\alpha_s)-2\ \sum_{i=1}^{r}k_i(\alpha_s,\alpha_i)+\sum_{j=1}^{r}\sum_{i=1}^{r}k_ik_j(\alpha_i,\alpha_j)$$

$$=\left(\alpha_s-\sum_{i=1}^{r}k_i\alpha_i,\alpha_s-\sum_{i=1}^{r}k_i\alpha_i\right)$$

$$=0$$

故有

$$\beta_s-\sum_{i=1}^{r}k_i\beta_i=0$$

即

$$\beta_s=\sum_{i=1}^{r}k_i\beta_i$$

则$\beta_1,\beta_2,\cdots,\beta_r$是向量组的极大无关组。

其次证明，将$\alpha_1,\alpha_2,\cdots,\alpha_r$标准正交化后得标准正交组$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r$。即

$$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r)=(\alpha_1,\alpha_2,\cdots,\alpha_r)\begin{pmatrix}t_{11}&t_{12}&\cdots t_{1r}\\0&t_{22}&\cdots t_{2r}\\\cdots&\cdots&\cdots\\0&0&\cdots t_{rr}\end{pmatrix}\qquad(1)$$

则

$$(\eta_1,\eta_2,\cdots,\eta_r)=(\beta_1,\beta_2,\cdots,\beta_r)\begin{pmatrix}t_{11} & t_{12} & \cdots t_{1r}\\ 0 & t_{22} & \cdots t_{2r}\\ \cdots & \cdots & \cdots\\ 0 & 0 & \cdots t_{rr}\end{pmatrix} \qquad (2)$$

也是由$\beta_1,\beta_2,\cdots,\beta_r$所得到的一个标准正交组。

事实上，因为

$$(\alpha_i,\alpha_j)=(\beta_i,\beta_j)$$

则有

$$\begin{aligned}(\eta_i,\eta_j)&=(t_{1i}\beta_1+t_{2i}\beta_2+\cdots+t_{ri}\beta_r,\ t_{1j}\beta_1+t_{2j}\beta_2+\cdots+t_{rj}\beta_r)\\ &=(t_{1i}\alpha_1+t_{2i}\alpha_2+\cdots+t_{ri}\alpha_r,\ \ t_{1j}\alpha_1+t_{2j}\alpha_2+\cdots+t_{rj}\alpha_r)\\ &=(\varepsilon_i,\ \varepsilon_j)\end{aligned}$$

故$\eta_1,\eta_2,\cdots$，η_r是一个标准正交组。

然后由本章练习三的第3 题，存在一个正交变换 σ，使得

$$\sigma(\varepsilon_i)=\eta_i \qquad i=1,2,\cdots,r \qquad (3)$$

最后证明

$$\sigma(\alpha_i)=\beta_i$$

事实上，由（1）式有

$$\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r)=(\sigma(\alpha_1),\sigma(\alpha_2),\cdots\sigma(\alpha_r))T \qquad (4)$$

再由（3）和（2）式有

$$\begin{aligned}\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_r)&=(\eta_1,\eta_2,\cdots,\eta_r)\\ &=(\beta_1,\beta_2,\cdots,\beta_r)T \qquad (5)\end{aligned}$$

比较（4），（5）则有

$$\sigma(\alpha_i)=\beta_i \quad i=1,2,\cdots,\ r$$

任取β_s，则有

$$\beta_s=\sum_{i=1}^{r}k_i\beta_i$$

由1）的证明有，

$$\alpha_s=\sum_{i=1}^{r}k_i\alpha_i$$

于是有

$$\sigma(\alpha_s)=\sum k_i\sigma(\alpha_i)=\sum k_i\beta_i=\beta_s$$

从而问题得证。

16．证：设

$$A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$$

其中$\alpha'_i=(a_{1i},a_{2i},\cdots,a_{ni})\qquad i=1,2,\cdots,n$。

下面证明必要性：设$x=(x_1',x'_2,\cdots,x'_n)$是（1）的任意解，则可有

$$\beta=x'_1\alpha'_1+x'_2\alpha_2'+\cdots+x'_n\alpha'_n$$

又设$x_0=(x^0_1,x^0_2\cdots,\ x^0_n)$是（2）的任一解向量，则

$$a_{1i}x^0_1+a_{2i}x^0_2+\cdots+a_{ni}x^0_n=0\qquad i=1,2,\cdots,n.$$

即

$$(\alpha_i,x_0)=0$$

故α_i的任一线性组合与x_0正交，即有

$$(\beta,x_0)=0$$

充分性：设W是$\alpha_1,\alpha_2,\cdots,\alpha_n$生成的子空间。

$$W=L(\alpha_1,\alpha_2,\cdots,\alpha_n)$$

则$W^{\perp}$恰好是（2）的解空间，故

$$\beta\perp W^{\perp}$$

于是

$$\beta\in W$$

故β是$\alpha_1,\alpha_2,\cdots,\alpha_n$的某一线性组合，即

$$\beta=x_1'\alpha_1'+x_2'\alpha_2'+\cdots+x'_n\alpha_n'$$

因此是（1）的解。

后　　记

本书是在我们讲授高等代数的讲稿基础上改编写成的。这次改编充分注意到广大自学读者的具体情况和实际需要。

袁秉成同志参加编写了第二和第三章；高绪珏同志参加编写了第四、第五和第六章；迟志敏同志参加编写了第八、第九和第十这三章。

由于我们的水平所限，书中的问题一定不少，希望大家提出批评、指正。

编　者

1981年11月

［Ｇｅｎｅｒａｌ　Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝高等代数　　（下册）
作者＝
页数＝６１６
ＳＳ号＝０
出版日期＝
Ｖｓｓ号＝６０１５４６９２

目　录